PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/54, 9/12, A61K 31/70

A2

(11) International Publication Number: WO 00/52173

(43) International Publication Date: 8 September 2000 (08.09.00)

(21) International Application Number: PCT/CA00/00223

(22) International Filing Date: 2 March 2000 (02.03.00)

(30) Priority Data:
60/122,516 2 March 1999 (02.03.99) US

(71) Applicant *(for all designated States except US)*: NPS ALLELIX CORP. [CA/CA]; 6850 Goreway Drive, Mississauga, Ontario L4V 1V7 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MUNROE, Donald [CA/CA]; 27 Wakefield Lane, Waterdown, Ontario L0R 2N3 (CA). GUPTA, Ashwani [CA/CA]; 7031 Dunrobin Way, Mississauga, Ontario L5N 6Y4 (CA). FALZONE, Germaine, R. [CA/CA]; 34 Brady Crescent, North York, Ontario M3L 2B5 (CA).

(74) Agents: VAN ZANT, Joan, M. et al.; 77 Bloor Street West, Suite 1407, Toronto, Ontario M5S 1M2 (CA).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: CLONED HUMAN SPHINGOSINE KINASE HOMOLOGUES

(57) Abstract

The present invention provides newly identified and isolated polynucleotides and their polypeptides and their uses and in particular to newly identified and isolated polynucleotides and polypeptides of the sphingosine kinase family. Three isolated polynucleotides and polypeptides for three human SK homologues are described: SKA, SKB and SKC.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Cloned Human Sphingosine Kinase Homologues

Field of Invention

The present invention is related to the field of molecular biology. In particular the present invention is related to newly identified and isolated polynucleotides and their polypeptides and their uses and in particular to newly identified and isolated polynucleotides and polypeptides of the sphingosine kinase family.

Background

Sphingolipids are complex structural lipids which are found in membranes. One of the more prominent sphinoglipids is sphingomyelin. Sphingomyelin is hydrolysized by sphinogmyelinase to form ceramide which in turn is metabloized by ceramidase to form sphingosine. Sphingosine kinase (SK) is the enzyme which phosphorylates sphingosine to form sphingosine 1-phosphate (S1P) and thereby SK in effect controls S1P production.

It has been suggested that S1P plays an intracellualar role in cell proliferation and inhibits and/or blocks apoptosis, among other things. S1P also acts as an endogenous ligand for certain G-protein coupled receptors, including the edg-3 receptor It has also been found that SK inhibitors block thrombin signalling pathways and induce apoptosis.

Accordingly, in view of the role of SK and SIP in key pathways, it would be desirable to have the cloned human SK homologues for use as drug targets. In particular, using the human SK homologues, SK inhibitors could be determined which inhibitors could be used in anti-proliferative diseases: cancer, psoriasis, reactive gliosis; or to surpress inappropriate cell survival and to block inflammatory S1P production in neurodegenerative dieseases, demyelinating diseases, asthma, and allergies

Summary of the Invention

The present invention provides the isolated polynucleotides and polypeptides for the human SK homologues. In particular, the present invention provides three isolated polynucleotides and polypeptides for the three human SK homologues: SKA; SKB; and SKC; and variants thereof.

In accordance with an aspect of the present invention there is provided isolated polypeptides for the human SK homologues SKA, SKB and SKC comprising the sequences as set out in Figures 3, 6 and 9, respectively, and variants thereof.

In accordance with another aspect of the present invention there is provided isolated polynucleotides of SKA, SKB and SKC comprising the sequences as illustrated in Figures 2, 5 and 8, respectively, and variants thereof.

In accordance with a further aspect of the present invention there are provided isolated polynucleotides encoding human SKA, SKB and SKC, including mRNAs, cDNAs, genomic DNAs. In addiition, embodiments of the invention include diagnostic, prophylactic, clinical or therapeutical useful variants of these isolated nucleotide sequences SKA, SKB and SKC and compositions thereof. Also included in an aspect of the invention are natuarally occurring allelic variants of SKA, SKB and SKC and polypeptides encoded thereby.

In accordance with another aspect of the invention, there are provided methods for producing the polypeptides for SKA, SKB and SKC and for determining inhibitors to such polypeptides, including antibodies.

In accordance with another aspect there are polynucleotides that hybridize to SKA, SKB and SKC nucleotide sequences, particularly under stringent conditions.

In accordance with yet another aspect of the invention, there are provided methods for identifying compounds which interact with the polypeptide or polynucleotide of SKA, SKB or SKC.

There are also provided compositions comprising a polypeptide or polynucleotide of SKA, SKB or SKC for administration to a cell or to a multicellular organism.

Description of the Figures

Figure 1 is an illustration of the full length nucleotide sequence of the cDNA of human SKA.

Figure 2 is an illustration of the nucleotide sequence of the coding region of SKA.

Figure 3 is an illustration of the predicted amino acid sequence of SKA as illustrated in Figure 2.

Figure 4 is an illustration of the full length nucleotide sequence of the cDNA of human SKB.

Figure 5 is an illustration of the nucleotide sequence of the coding region of SKB.

Figure 6 is an illustration of the predicted amino acid sequence of SKB as illustrated in Figure 5.

Figure 7 is an illustration of the full length nucleotide sequence of the cDNA of human SKC.

Figure 8 is an illustration of the nucleotide sequence of the coding region of SKC.

Figure 9 is an illustration of the predicted amino acid sequence of SKC as illustrated in Figure 8.

Figure 10 is an illustration of the alignment of the amino acid sequences of human SKA, SKB and SKC.

Figure 11 is an illustration of the results of the phosphorylation assays exemplified in Example 4 for SKA, SKB and SKB.

Detailed Description

Definitions

The following definitions are used herein for the purpose of describing particular terms used in the application. Any terms not specifically defined should be given the meaning commonly understood by one of ordinary skill in the art to which the invention pertains.

"Biologically Active" refers to those forms, fragments, or domains of any sphingosine kinase polypeptide which retain at least some of the biological and/or antigenic activities of a naturally occurring sphingosine kinase.

"Chimeric" molecules may be constructed by introducing all or part of the nucleotide sequence of this invention into a vector containing additional nucleic acid sequence which might be expected to change any one (or more than one) of the following characteristics: cellular location, distribution, ligand-binding affinities, interchain affinities, degradation/turnover rate, signaling, etc.

"Derivative" refers to those amino acid sequences and nucleotide sequences which have been chemically modified. Such techniques for polypeptide derivatives include: ubiquitination; labeling (see above); pegylation (derivatization with polyethylene glycol); and chemical insertion or substitution of amino acids such as ornithine which do not normally occur in human proteins. A nucleotide sequence derivative would encode an amino acid which retains its essential biological activity and characteristics of the natural molecule.

As used herein "human sphingosine kinase" refers to the isolated polypeptide or polynucleotide sequences of the different isoforms of human sphingosine kinase, including human SKA, human SKB and human SKC, in either naturally occurring or synthetic form.

As used herein, "human sphingosine kinase A" or "human SKA" refers to the polynucleotide or polypeptide of an isoform of human sphingosine kinase as illustrated by the sequences of Figure 2 and 3, respectively, and by polypeptide sequences which preferably have at least 85% sequence identity with each other and Figure 3, and more preferably at least 90% sequence identity with each other and Figure 3, and most preferably at least 95% sequence identity with each other and Figure 3, or polynucleotide sequences which encode such polypeptide sequence identities.

As used herein, "human sphingosine kinase B" or "human SKB" refers to the polyucleotide or polypeptide of an isoform of human sphingosine kinase as illustrated by the sequences of Figure 5 and 6, respectively, and to the polypeptide sequences which preferably have at least 85% sequence identity with each other and Figure 6, and more preferably at least 90% sequence identity with each other and Figure 6, and most preferably at least 95% sequence identity with each other and Figure 6 or polynucleotide sequences which encode such polypeptide sequence identities.

As used herein, "human sphingosine kinase C" or "human SKC" refers to the polynucleotide or polypeptide of an isoform of human sphingosine kinase as illustrated by the sequences of Figure 8 and 9, respectively, and by the polypeptide sequences which preferably have at least 85% sequence identity with each other and Figure 9, and more preferably at least 90% sequence identity with each other and Figure 9, and most preferably at least 95% sequence identity with each other and Figure 9 or polynucleotide sequences which encode such polypeptide sequence identities.

"Inhibitor" is any substance which retards or prevents a biochemical, cellular or physiological reaction or response. Common inhibitors include but are not limited to antisense molecules, antibodies, and antagonists.

"Insertions" or "deletions" are typically in the range of about 1 to 5 amino acids and do not result in a change in biological activity of the polypeptide. The variation allowed may be experimentally determined by producing the peptide synthetically or by systematically making insertions, deletions, or substitutions of nucleotides in the human sphingosine kinase sequence using recombinant DNA techniques.

As used herein "isolated" means separated from nucleotide sequences s that encode other proteins or from other peptides. For example, a polypeptide or polynucleotide naturally present in a living organism is not "isolated" but when separated from the coexisting nucleotides/peptides it is "isolated".

"Nucleotide sequences" as used herein are oligonucleotides, polynucleotides, and fragments or portions thereof, and are DNA or RNA of genomic or synthetic origin which may be single or double stranded, and represent the sense or complement or antisense strands.

An "oligonucleotide" is a stretch of nucleotide residues, which has a

sufficient number of bases to be used as an oligomer, amplimer or probe in a polymerase chain reaction (PCR). Oligonucleotides are prepared from genomic or cDNA sequence and are used to amplify, reveal or confirm the presence of a similar DNA or RNA in a particular cell or tissue. Oligonucleotides or oligomers comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 35 nucleotides, preferably about 25 nucleotides.

An "oligopeptide" is a short stretch of amino acid residues and may be expressed from an oligonucleotide. It may be functionally equivalent to and the same length as (or considerably shorter than) a "fragment", "portion", or "segment" of a polypeptide. Such sequences comprise a stretch of amino acid residues of at least about 5 amino acids and often about 17 or more amino acids, typically at least about 9 to 13 amino acids, and of sufficient length to display biological and/or antigenic activity.

As used herein "purified" refers to amino acid sequences that are removed from their natural environment, and are isolated or separated, and are at least 60% free, preferably at least 75 % free, and most preferably at least 90% free from other components with which they are naturally associated.

A "portion" or "fragment" of a nucleotide or nucleic acid sequence comprises all or any part of the sequence having fewer nucleotides than about 6 kb, preferably fewer than about 1 kb. A portion or fragment can be used as a probe. Such probes may be labeled with reporter molecules using nick translation, Klenow fill-in reaction, PCR or other methods well known in the art. To optimize reaction conditions and to eliminate false positives, nucleic acid probes may be used in Southern, Northern or in situ hybridizations to determine whether DNA or RNA encoding spingosine kinase is present in a cell type, tissue, or organ.

"Probes" may be derived from naturally occurring, recombinant, or chemically synthesized single - or double - stranded nucleic acids or be chemically

synthesized. They are useful in detecting the presence of identical or similar sequences.

"Reporter" molecules are those radionuclides , enzymes, fluorescent, chemiluminescent, or chromogenic agents which associate with, establish the presence of, and may allow quantification of a particular nucleotide or amino acid sequence.

A "signal or leader sequence" can be used, when desired, to direct the polypeptide through a membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous sources by recombinant DNA techniques.

Amino acid "substitutions" are conservative in nature when they result from replacing one amino acid with another having similar structural and/or chemical properties, such as the replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

"Standard" is a quantitative or qualitative measurement for comparison. It is based on a statistically appropriate number of normal samples and is created to use as a basis of comparison when performing diagnostic assays, running clinical trials, or following patient treatment profiles.

"Stringent conditions" is used herein to mean conditions that allow for hybridization of substantially related nucleic acid sequences. . Such hybridization conditions are described by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989. Generally, stringency occurs within a range from about 5 °C below the melting temperature of the probe to about 20 °C – 25 °C below the melting temperature. As understood by ordinary skilled persons in the art, the stringency conditions may be altered in order to identify or detect identical or related nucleotide sequences. Factors such as the length and nature

(DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.) and the concentration of the salts and other componenets (e.g. the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency.

"Sequence Identity" is known in the art, and is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences, particularly, as determined by the match between strings of such sequences. Sequence identity can be readily calculated by known methods (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two sequences, the term is well known to skilled artisans (see, for example, *Sequence Analysis in Molecular Biology*; *Sequence Analysis Primer*; and Carillo, H., and Lipman, D., SIAM *J. Applied Math.*, **48**: 1073 (1988)). Methods commonly employed to determine identity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., SIAM *J. Applied Math.*, **48**: 1073 (1988) or, preferably, in Needleman and Wunsch, *J. Mol. Biol.*, **48**: 443-445, 1970, wherein the parameters are as set in version 2 of DNASIS (Hitachi Software Engineering Co., San Bruno, CA). Computer programs for determining identity are publicly available. Preferred computer program methods to determine identity between two sequences include, but are not limited to, GCG program package (Devereux, J., *et al.*, *Nucleic Acids Research* **12(1)**: 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. *et al.*, *J. Molec. Biol.* **215**: 403-410 (1990)). The BLASTX program is publicly available from NCBI (blast@ncbi.nlm.nih.gov) and other sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, MD

20894; Altschul, S., *et al.*, *J. Mol. Bio.* **215**: 403-410 (1990)).Computational Molecular Biology, Lesk, A.M, ed. Unless specified otherwise in the claims, the percent identity for the purpose of interpreting the claims shall be calculated by the Needleman and Wucnsch algorithm with the parameters set in version 2 of DNASIS.

"Variants" are polynucleotides or polypeptides that differ from a reference polynucleotide or polypeptide, respectively, but retain essential properties of the reference, preferably, in the case of polypeptides the variant retains the biological activity of the naturally occurring polypeptide . A typical variant of a polynucleotide differs in nucleotide sequence from another reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequences of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, insertions and deletions in the polypeptide encoded by the reference sequences, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and ,in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, insertions and deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurrring variants of polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

Description

The invention relates to novel polypeptides and polynucleotides for human sphingosine kinase as described in greater detail below. The invention particularly relates to the three sphingosine kinase homologues: human SKA, human SKB and human SKC. More particularly, human SKA, SKB and SKC having the nucleotide

sequences as set out in Figures 2, 5 and 8 for SKA, SKB and SKC, respectively, and variants thereof are provided for herein.

In addition, the polypeptides of the invention include the polypeptides comprising the sequences as set out in Figures 3, 6 and 9 as well as variants of these polypeptides, particularly variants which retain the biological activity of the naturally occurring sphingosine kinase.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full length polypeptides of the invention.

The polynucleotides comprising sequences encoding human SKA, SKB and SKC (or their complement) and variants thereof have numerous applications in techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use in the construction of oligomers for PCR, use for chromosome and gene mapping, use in the recombinant production of SKA, SKB and SKC, and use in generation of antisense DNA or RNA, their chemical analogs and the like. Uses of nucleotides encoding SKA, SKB and SKC disclosed herein are exemplary of known techniques and are not intended to limit their use in any technique known to a person of ordinary skill in the art. Furthermore, the nucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, e.g., the triplet genetic code, specific base pair interactions, etc.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of human SKA, SKB and SKC encoding nucleotide sequences may be produced. Some of these will only bear minimal homology to the nucleotide sequence of the known and naturally occurring SKA,

SKB and SKC. The invention has specifically contemplated each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring SKA, SKB and SKC, and all such variations are to be considered as being specifically disclosed.

Although the nucleotide sequences which encode SKA, SKB and SKC, their derivatives or their variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring SKA, SKB and SKC, respectively, under stringent conditions, it may be advantageous to produce nucleotide sequences encoding SKA, SKB and SKC or its derivatives possessing a substantially different codon usage. Codons can be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic expression host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SKA, SKB and SKC and/or its derivatives without altering the encoded amino acid sequence include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

Human genes often show considerable actual polymorphism; that is, variation in nucleotide sequence among a fraction of the entire human population. In many cases this polymorphism can result in one or more amino acid substitutions. While some of these substitutions show no demonstrable change in function of the protein, others may produce varying degrees of functional effects. In fact, many natural or artificially produced mutations can lead to expressible human SK proteins. Each of these variants, whether naturally or artificially produced, is considered to be equivalent and specifically incorporated into the present invention.

Nucleotide sequences encoding human SKA, SKB and SKC may be joined to a variety of other nucleotide sequences by means of well established recombinant

DNA techniques (Sambrook J et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor NY; or Ausubel FM et al (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York City). Useful nucleotide sequences for joining to human SK include an assortment of cloning vectors such as plasmids, cosmids, lambda phage derivatives, phagemids, and the like. Vectors of interest include expression vectors, replication vectors, probe generation vectors, sequencing vectors, etc. In general, vectors of interest may contain an origin of replication functional in at least one organism, convenient restriction endonuclease sensitive sites, and selectable markers for one or more host cell systems.

Human SK specific hybridization probes are capable of hybridizing with naturally occurring nucleotide sequences encoding human SKA, SKB and SKC. Such probes may also be used for the detection of similar sequences and should preferably contain at least 60% nucleotide identity to SK sequence. The hybridization probes of human SK may be derived from the nucleotide sequence presented in the Figures for the full length sequence for SKA, SKB and SKC, namely, Figures 1, 4 and 7, respectively, or from genomic sequences including promoter, enhancers, introns or 3'-untranslated regions of the native gene. Hybridization probes may be labeled by a variety of reporter molecules using techniques well known in the art. Preferably, the hybridization probes incorporate at least 15 nucleotides, and preferably at least 25 nucleotides, of the SK protein. Suitable hybridization probes would include: consensus fragments, for example, those regions of the human SK isoforms that are identical, as particularly exemplified in Figure 10.

It will be recognized that many deletional or mutational analogs of nucleic acid sequences for human SK will be effective hybridization probes for human SK nucleic acid. Accordingly, the invention relates to nucleic acid sequences that hybridize with such SK encoding nucleic acid sequences under stringent conditions.

Stringent conditions will generally allow hybridization of sequence with at

least about 70% sequence identity, more preferably at least about 80-85% sequence identity, even more preferably at least about 90% sequence identity, and most preferably with at least about 95% sequence identity. Hybridization conditions and probes can be adjusted in well-characterized ways to achieve selective hybridization of human-derived probes. Nucleic acid molecules that will hybridize to human SK encoding nucleic acid under stringent conditions can be identified functionally, using methods outlined above, or by using for example the hybridization rules reviewed in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989. Without limitation, examples of the uses for hybridization probes include: histochemical uses such as identifying tissues that express human SK; measuring mRNA levels, for instance to identify a sample's tissue type or to identify cells that express abnormal levels of human SK; and detecting polymorphisms in the human SK. RNA hybridization procedures are described in Maniatis et al. Molecular Cloning, a Laboratory Manual (Cold Spring Harbor Press, 1989). PCR as described US Patent No's. 4,683,195; 4,800,195; and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequence which encodes the human SK sequences of the invention. Such probes used in PCR may be of recombinant origin, chemically synthesized, or a mixture of both. Oligomers may comprise discrete nucleotide sequences employed under optimized conditions for identification of human SK in specific tissues or diagnostic use. The same two oligomers, a nested set of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for identification of closely related DNA's or RNA's. Rules for designing PCR primers are now established, as reviewed by PCR Protocols, Cold Spring Harbor Press, 1991. Degenerate primers, i.e., preparations of primers that are heterogeneous at given sequence locations, can be designed to amplify nucleic acid sequences that are highly homologous to, but not identical to human SK. Strategies are now available that allow for only one of the primers to be required to specifically hybridize with a known sequence. See, Froman et al., Proc. Natl. Acad. Sci. USA 85: 8998, 1988 and Loh et al., Science 243: 217, 1989. For example, appropriate nucleic acid primers can be ligated to the nucleic acid sought to be amplified to provide the hybridization partner for one of the primers. In this way, only one of the primers

need be based on the sequence of the nucleic acid sought to be amplified. PCR methods of amplifying nucleic acid will utilize at least two primers. One of these primers will be capable of hybridizing to a first strand of the nucleic acid to be amplified and of priming enzyme-driven nucleic acid synthesis in a first direction. The other will be capable of hybridizing the reciprocal sequence of the first strand (if the sequence to be amplified is single stranded, this sequence will initially be hypothetical, but will be synthesized in the first amplification cycle) and of priming nucleic acid synthesis from that strand in the direction opposite the first direction and towards the site of hybridization for the first primer. Conditions for conducting such amplifications, particularly under preferred stringent hybridization conditions, are well known. See, for example, PCR Protocols, Cold Spring Harbor Press, 1991.

Other means of producing specific hybridization probes for human SK include the cloning of nucleic acid sequences encoding human SK or human SKvariants or derivaitves into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate reporter molecules.

It is possible to produce a DNA sequence, or portions thereof, entirely by synthetic chemistry. After synthesis, the nucleic acid sequence can be inserted into any of the many available DNA vectors and their respective host cells using techniques which are well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into the nucleotide sequence. Alternately, a portion of sequence in which a mutation is desired can be synthesized and recombined with longer portion of an existing genomic or recombinant sequence.

The nucleotide sequence for human SK can be used in an assay to detect inflammation or disease associated with abnormal levels of SK expression. The cDNA can be labeled by methods known in the art, added to a fluid, cell or tissue sample from a patient, and incubated under hybridizing conditions. After an

incubation period, the sample is washed with a compatible fluid which optionally contains a reporter molecule. After the compatible fluid is rinsed off, the reporter molecule is quantitated and compared with a standard as previously defined.

A diagnostic test for aberrant expression of SK can accelerate diagnosis and proper treatment of abnormal conditions of SK activity.

New nucleotide sequences can be assigned to chromosomal subregions by physical mapping. The mapping of new genes or nucleotide sequences provide useful landmarks for investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome, such as ataxia telangiectasia (AT), has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 1 1q22-23 (Gatti et al (1988) Nature 336:577-580), any sequences mapping to that area may represent or reveal genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in gene sequence between normal and carrier or affected individuals.

Nucleotide sequences encoding human SK may be used to produce a purified oligo - or polypeptide using well known methods of recombinant DNA technology. Goeddel (1990, Gene Expression Technology, Methods and Enzymology, Vol. 185, Academic Press, San Diego CA) is one among many publications which teach expression of an isolated nucleotide sequence. The oligopeptide may be expressed in a variety of host cells, either prokaryotic or eukaryotic. Host cells may be from the same species from which the nucleotide sequence was derived or from a different species. Advantages of producing an oligonucleotide by recombinant DNA technology include obtaining adequate amounts of the protein for purification and the availability of simplified purification procedures.

Cells transformed with DNA encoding human SK may be cultured under conditions suitable for the expression of human kinases and recovery of such peptides

from cell culture. Human SK produced by a recombinant cell may be secreted or may be contained intracellularly, depending on the particular genetic construction used. In general, it is more convenient to prepare recombinant proteins in secreted form. Purification steps vary with the production process and the particular protein produced. Often an oligopeptide can be produced from a chimeric nucleotide sequence. This is accomplished by ligating the nucleotides from human SK or a desired portion of the polypeptide to a nucleic acid sequence encoding a polypeptide domain which will facilitate protein purification (Kroll DJ et al (1993) DNA Cell Biol. 12:441-53).

In addition to recombinant production, fragments of human SK may be produced by direct peptide synthesis using solid-phase techniques (e.g. Stewart at al (1969) Solid-Phase Peptide Synthesis, WH Freeman Co., San Francisco QA; Merrifield J (1963) J Am Chem. Soc. 85:2149-2154). Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Foster City, CA) in accordance with the instructions provided by the manufacturer. Additionally, a particular portion of human SK may be mutated during direct synthesis and combined with other parts of the peptide using chemical methods.

Human SK for antibody induction does not require biological activity: however, the protein must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids. They should mimic a portion of the amino acid sequence of the protein and may contain the entire amino acid sequence. An antigenic portion of human SK may be fused to another protein such as keyhole limpet hemocyanin, and the chimeric molecule used for antibody production.

Antibodies specific for human SK may be produced by inoculation of an appropriate animal with the polypeptide or an antigenic fragment. An antibody is specific for human SK if it is produced against an epitope of the polypeptide and binds to at least part of the natural or recombinant protein. Antibody production

includes not only the stimulation of an immune response by injection into animals, but also analogous processes such as the production of synthetic antibodies, the screening of recombinant immunoglobulin libraries for specific- binding molecules (e.g. Orlandi R et al (1989) PNAS 86:3833-3837, or Huse WD et al (1989) Science 256:1275-1281) or the in vitro stimulation of lymphocyte populations. Current technology (Winter G and Mistein C (1991) Nature 349:293-299) provides for a number of highly specific binding reagents based on the principles of antibody formation. These techniques may be adapted to produce molecules which specifically bind SK.

An additional embodiment of the subject invention is the use of human SK specific antibodies, inhibitors, ligands or their analogs as bioactive agents to treat inflammation or disease possibly including, but not limited to viral, bacterial or fungal infections; allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or other conditions which activate the genes of kidney, lung, heart, lymphoid or tissues of the nervous system.

Bioactive compositions comprising agonists, antagonists, receptors or inhibitors of human SK may be administered in a suitable therapeutic dose determined by any of several methodologies including clinical studies on mammalian species to determine maximal tolerable dose and on normal human subjects to determine safe dose. Additionally, the bioactive agent may be complexed with a variety of well established compounds or compositions which enhance stability or pharmacological properties such as half-life. It is contemplated that the therapeutic, bioactive composition may be delivered by intravenous infusion into the bloodstream or any other effective means which could be used for treating problems involving aberrant expression of the EDG-7 gene.

The examples below are provided to illustrate the subject invention. These examples are provided by way of illustration and are not included for the purpose of limiting the invention.

Example 1 Cloning of PSKA

A. Diagnostic PCR of various templates for the presence of human SKA cDNA

The following pairs of primers were designed:

5' end Primers

SK1F 5' AAC CCG CGC GGC GCA AGG GCA AGG C 3'

SK2F 5' AAG GGC AAG GCC TTG CAG CTC TTC C 3'

3' end Primers

SK1R 5' CAG GCC GCT CCA TGA GCC CGT TCA C 3'

SK2R 5' GCA TCA GCC CGT CTC CAG ACA TGA 3'

Using these primers, PCR was conducted under the conditions as set out below on the templates from the following sources:

Template source DNA of cDNA Libraries prepared from Human Lung Fibroblasts WI-38 (Origene Technologies Inc., Cat. No. DLH-102), Human Liver (Origene Technologies Inc., Cat. No. DLH-100), cultured human Jurkat T-cells (Origene Technologies Inc., Cat. No. DLH-115), HeLa cultured cells (Origene Technologies Inc., Cat. No. DLH-103), Human kidney proximal tubules (ATCC), HeLa cultured cells (Invitrogen, Cat. No. A550-26), Human Lung (Clonetech, Cat. No.7114-1), HeLa cultured cells (Clonetech, Cat. No. HL5013a), human small intestine (Clonetech, Cat. No.HL1133a). Each template was amplified with each pair of primers under the following condition of PCR amplification by using Expand TM PCR kit of Boehringer Mannheim (Catalogue no. 1681-842).

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3

0.4 µl of 25mM dNTP mix

0.6 µl of Primer SKF1 or SKF2 (10pm/µl)

0.6 µl of Primer SKR1 or SKR2 (10pm/µl)

0.3µl of Enzyme (3unit)

15.1 µl water

1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 1 min
	62°C for 1 min
	68°C for 30 sec
Incubate:	68°C for 8 min
Hold:	4°C

A 200 bp (approximately) DNA fragment was amplified from all templates except Origene's HeLa cDNA library and Clonetech's small intestine cDNA library. The cDNA library from cultured HeLa cells (Invitrogen) appeared to contain PSKA clones.

B. PCR Screening of HeLa cDNA Library

10 000 (10K) clone pools from cDNA library form cultured HeLa cells (Invitrogen, Cat. No. A550-26): Approximately, ten thousand clones were grown on an agar plate, scraped and re-suspended in one ml of 2X YT + 20% glycerol. Overall, 610 pools (10K) were prepared. Equal proportions of twelve 10K pools were mixed to prepare 120K pools. In all, there were fifty one 120K pools. All pools are kept as

frozen stocks at 80°C. For PCR screening, a small portion of frozen stock was re-suspended in 100 ul of 2X YT + 20% glycerol and used as template.

C. Screening of 120K and 10K bacterial

All fifty one 120K bacterial pools and 10K pools of positive 120K pools were amplified under the following condition of PCR amplification by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

SK1F	5' AAC CCG CGC GGC GCA AGG GCA AGG C
SK1R	5' CAG GCC GCT CCA TGA GCC CGT TCA C 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer SKF1 (10pm/µl)
0.6 µl of Primer SKR1 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 40 sec
	60°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min
Hold:	4°C

Majority of 120K bacterial pools was found positive indicating that PSKA is an abundantly expressed gene. Four 10 K pools (62, 64, 74, 403, and 404 from selected positive 120K pools) were found positive.

D. PCR screening of sub-pools of 10K pool #403:

The bacterial colonies were grown from the positive 10K pool #403 on the agar plate. Plugs containing 300 – 1000 bacterial colonies were lifted from the agar plate. The bacterial colonies were re-suspended into 500 ul of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template to amplify with the following pair of primers.

SK1F	5' AAC CCG CGC GGC GCA AGG GCA AGG C
SK1R	5' CAG GCC GCT CCA TGA GCC CGT TCA C 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer SKF1 (10pm/µl)
0.6 µl of Primer SKR1 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 40 sec
	60°C for 40 sec
	68°C for 40 sec

Incubate: 68°C for 8 min

Hold: 4°C

Several positive sub-pools were identified. Sub-pool #403-42 was used further to isolate PSKA clone.

E. Hybridization screening of sub-pool # 403-42

Bacterial colonies were grown from sub-pool #403-42 on the agar plate and were transferred to nylon filters. Filter hybridization was carried out using 200 bp DNA fragment (amplified by SK1F and SK1R from template 10K pool no. 62) as probe. The following hybridization conditions were employed:

5X SSPE
5X Denharts solution (1% Ficoll, 1% Polyvinylpyrrolidone, 1% BSA)
25µg/ml fish sperm DNA

Hybridise at 65°C overnight.

The filters were washed 2 times in 2X SSPE and 0.1% SDS at room temperature for 30 minutes each, then 2 times in 2XSSPE and 0.1% SDS at 50°C for 20 minutes each and finally two times in 0.1XSSPE and 0.1% SDS.

One positive plug 403-42.1 was identified.

F. PCR screening of plug #403-42.1

Bacterial colonies from plug #403-42.1 were grown on the agar plate, picked and re-suspended into 100 ul of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template to amplify with the following pair of primers.

SK1F 5' AAC CCG CGC GGC GCA AGG GCA AGG C
SK1R 5' CAG GCC GCT CCA TGA GCC CGT TCA C 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer SKF1 (10pm/µl)
0.6 µl of Primer SKR1 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 40 sec
	60°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min
Hold:	4°C

A single isolated bacterial colony (403-42.1-P1C6-P1C3) was identified as a positive colony to contain PSKA cDNA. The plasmid and clone was given ID as pc3-PSKA#403-1. The plasmid DNA was prepared using midi-plasmid preparation kit (Qiagen, catalogue no. 12245) to use for sequencing and transfections.

Example 2 Cloning of SKB

A. Screening of 120K bacterial pools and 10K bacterial

All fifty one 120K bacterial pools were amplified under the following condition of PCR amplification by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

PSKB-C4F2	5' TGC AAA TCT CTA GAA GAT GAC GGT G 3'
PSKB-C4R3	5' TAT ACT CAA ACT ACT GGT CTC TCC AAG 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer PSKB-C4F2 (10pm/µl)
0.6 µl of Primer PSKB-C4R3 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 1 min
	62°C for 1 min
	68°C for 1 min
Incubate:	68°C for 8 min
Hold:	4°C

A DNA fragment of approximately 500bp was amplified from four 120K bacterial pools.

B. Screening of 10K bacterial pools

10K pools of positive three 120K pools were amplified under the following condition of PCR amplification by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

PSKB-C4F2	5' TGC AAA TCT CTA GAA GAT GAC GGT G 3'
PSKB-C4R3	5' TAT ACT CAA ACT ACT GGT CTC TCC AAG 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer PSKB-C4F2 (10pm/µl)
0.6 µl of Primer PSKB-C4R3 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
32 cycles:	94°C for 1 min
	58°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min
Hold:	4°C

Two 10K bacterial pools (308 and 532) were found positive.

C. Isolation of PSKB clone from 10K bacterial pool #532 by PCR Screening

The following three steps and three rounds of PCR were used to isolate individual positive clone of PSKB cDNA from 10K bacterial pool #532.

1. 10K bacterial pool #532 was plated on agar plates. 100-500 colonies were scraped in sub-pool and re-suspended in 100 µl of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template for PCR screening.
2. The positive sub-pool of 100-500 bacterial colonies was plated on agar plates. 20-50 colonies were scraped in sub-pool and re-suspended in 100 µl of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template for PCR screening.
3. The positive sub-pool of 20-50 bacterial colonies was plated on agar plates. The individual bacterial colonies were scraped and re-suspended in 100 µl of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template for PCR screening.

The PCR screening was done by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

PSKB-C4F2	5' TGC AAA TCT CTA GAA GAT GAC GGT G 3'
PSKB-C4R3	5' TAT ACT CAA ACT ACT GGT CTC TCC AAG 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer PSKB-C4F2 (10pm/µl)
0.6 µl of Primer PSKB-C4R3 (10pm/µl)

0.3 µl of Enzyme (3unit)

15.1 µl water

1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 1 min
	58°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min
Hold:	4°C

Two colonies (532 - P1A9 - P1G1 – P1E4 and 532 - P1A9 - P1G1 – P1E9) were found positive. They were given ID of pc3-PSKB#532-1 and pc3-PSKB#532-2. The plasmid DNA was prepared using mini-plasmid preparation kit (Qiagen, catalogue no. 12245) to use for sequencing and transfections.

Example 3 Cloning of SKC

A. Screening of 120K bacterial pools and 10K bacterial

All fifty one 120K bacterial pools were amplified under the following condition of PCR amplification by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

PSKC-F2	5' TTA ACA TAG ACA AAT ACG ACG GCA TCG 3'
PSKC-R1	5' ACA CAT CCA TGG CCA GCG AGT CC 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer PSKC-F2 (10pm/µl)
0.6 µl of Primer PSKC-R1 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 40 sec
	58°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min
Hold:	4°C

A DNA fragment of approximately 250bp was amplified from ten 120K bacterial pools.

B. Screening of 10K bacterial pools

10K pools of the four positive 120K pools were amplified under the following condition of PCR amplification by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

PSKC-F2	5' TTA ACA TAG ACA AAT ACG ACG GCA TCG 3'
PSKC-R1	5' ACA CAT CCA TGG CCA GCG AGT CC 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer PSKC-F2 (10pm/µl)
0.6 µl of Primer PSKC-R1 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 40 sec
	58°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min
Hold:	4°C

Three 10K bacterial pools (64, 320 and 330) were found positive.

C. Isolation of PSKC clone from 10K bacterial pool #330 by PCR Screening

The following three steps and three rounds of PCR were used to isolate individual positive clone of PSKC cDNA from 10K bacterial pool #330.

4. 10K bacterial pool #330 was plated on agar plates. 100-500 colonies were scraped in sub-pool and re-suspended in 100 µl of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template for PCR screening.

5. The positive sub-pool of 100-500 bacterial colonies was plated on agar plates. 20-50 colonies were scraped in sub-pool and re-suspended in 100 µl of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template for PCR screening.
6. The positive sub-pool of 20-50 bacterial colonies was plated on agar plates. The individual bacterial colonies were scraped and re-suspended in 100 µl of 2XYT + 20% glycerol. The bacterial re-suspensions were used as template for PCR screening.

The PCR screening were done by using Expand TM PCR system from Boehringer Mannheim (Catalogue no. 1681-842) with the following pair of primers.

PSKC-F2	5' TTA ACA TAG ACA AAT ACG ACG GCA TCG 3'
PSKC-R1	5' ACA CAT CCA TGG CCA GCG AGT CC 3'

Each reaction contained the following reagents:

2 µl of 10x PCR Buffer 3
0.4 µl of 25mM dNTP mix
0.6 µl of Primer PSKC-F2 (10pm/µl)
0.6 µl of Primer PSKC-R1 (10pm/µl)
0.3 µl of Enzyme (3unit)
15.1 µl water
1 µl DNA

PCR conditions:

Incubate:	94°C for 2 min
30 cycles:	94°C for 40 sec
	58°C for 40 sec
	68°C for 40 sec
Incubate:	68°C for 8 min

Hold: 4°C

Two colonies (330 – P1G3 – P1B8 – P2A9 and 330 – P1G3 – P4E10– P1B12) were found positive. They were given ID of pc3-PSKC#330-1 and pc3-PSKC#330-2. The plasmid DNA was prepared using mini-plasmid preparation kit (Qiagen, catalogue no. 12245) to use for sequencing and transfections.

Example 4 Phosphorylation activity of Human Sphingosine Kinase

A) Protocol of Phosphorylation Assay using Swiss 3T3 and 293-EBNA cells—Phosphate label

1) Swiss 3T3 cells were washed with PBS and harvested by scraping in 1 ml of protein buffer [0.1 M Tris-HCl, pH 7.4 containing 20% glycerol (v/v), 1mM mercaptoethanol, 1 mM EDTA, 1 mM Na_3VO_4 (Sigma, cat # S6508), 15 mM NaF, 10μg/ml leupeptin (Sigma, cat # L2023) and aprotinin (Sigma, cat# A6279), 1 mM PMSF and 0.5 mM 4-deoxypyridoxine(Sigma, cat # D0501)] (as described in Edsall et al., 1997).
2) Methods were taken from Edsall et al., 1997, with the following exceptions: once cells were lysed by freeze-thawing three times, the cytosolic fraction was prepared by centrifugation at 13,000 x g for 20 min at 4°C.
3) The phosphorylation reaction included, 80 μl of cytosolic fraction with 10 μl of sphingosine (Calbiochem, cat # 219535-S) (100 μM dissolved in a 4mg/ml solution of BSA). The reactions were started by adding 10 μl of [γ ^{33}P]-ATP (10-20 μCi, 10 mM) (Amersham cat # AH 9968) prepared in 100 mM $MgCl_2$. Samples were incubated at 37°C for one hour.
4) Lipids were extracted with 800 μl chloroform:methanol:concentrated HCl (100:200:1). Samples were vortexed and phases were separated by adding 240 μl

of chloroform and 240 µl of 2M KCl (as described in Olivera et al., 1994). Samples were vortexed and centrifuged at 11,000 x g for 5 min.

5) Lipids found in the organic phase were spotted on a TLC silica gel plate and run in the following solvent system: chloroform:methanol:acetic acid:water (60:30:5:5). A sphingosine 1-phsophate (S 1-P) (Sigma, cat #S-9666) and sphingosine (Sph) (standard was run alongside all experimental samples as well as a reaction tube containing no sphingosine as a negative control.

6) The sphingosine 1-phosphate and sphingosine standards were visualized using a $KMnO_4$ stain (100 ml H_2O: 4g $KMnO_4$, 4g $NaHCO_3$). TLC plates were exposed to a phosphor screen overnight.

7) S 1-P bands were visualized and quanitified with the Storm Phosphoimager (Molecular Dynamics, Sunnyvale, CA).

B) Transient transfection protocol for 293-EBNA

Day 1.

1) 100 mm plates of 293-EBNA with a confluency of 50-80% were used for transfection.

2) *SKA*, *SKB* , *SKC* and pcDNA3 (4 µg) DNA samples were diluted in 750 µl of DMEM/F12 (serum-free media) and 20 µl Plus Reagent (Lipofectamine Plus Kit, Life Technologies Cat. 10964-013), and incubated at room temperature for 15 min.

3) 30 µl Lipofectamine Reagent (Lipofectamine Plus Kit) was diluted in 750 µl DMEM/F12. The diluted Lipofectamine was then combined with the DNA/Plus mixture and incubated at RT for 15 min.

4) The 293-EBNA plates were washed once with PBS, and 5 ml DMEM/F12 was added to each plate.

5) DNA/Plus/Lipofectamine mixture was added to each plate of 293-EBNA cells. The plates were left for 3 hr at 37°C in a 5% CO_2 incubator.

6) The transfection medium was replaced with DMEM/F12 containing 10% FBS to recover overnight.

Day 3.

Media was removed and stored at -80°C. Cells were washed once with PBS and were harvested by scrapping cells in 1 ml of protein buffer (described above). Cells were lysed by freeze-thawing three times. Cytosolic fractions were obtained by centrifugation at 13,000 x g for 20 min at 4°C. Pellets of cell debris resulting from this spin was stored at -80°C for later use. Cytosolic protein preparations were stored at -80°C for future use.

C) Phosphorylation Assay using Three Fractions of SKA, SKB and SKC.

i) Cytosolic Fraction.

The phosphorylation assay was performed as outlined above, except 5 µl of Triton X-100 was added to each reaction tube.

ii) Debris Fraction

Each pellet consisting of cellular debris was re-suspended in 80 µl of protein buffer. The suspension was sonicated. This prep was used for the phosphorylation assay which was as outlined above in the cytosolic fraction.

iii) Cell Media

The cell media was dried down in a refrigerated speed-vac. The pellet was re-suspended in 80 µl of protein buffer. This preparation was used for the phosphorylation assay which was as outlined above in the cytosolic fraction.

D) Protocol of Phosphorylation Assay using 293-EBNA cells—Compound Label

Enzyme preparations of SKA and pcDNA3 were used in phosphorylation assays. The protocol was the same as mentioned above in A.3-7 with a few exceptions: i) 5 µl of Triton X-100 was added to each reaction tube and ii) ^{33}P-ATP was not used, only "cold" ATP was used in each reaction (10 mM in 100 mM $MgCl_2$) .

The results for the above phosphorylation assays indicated that SKA was involved in phosphorylating sphingosine to form sphingosine 1-phosphate. In particular, as indicated in Figure 11, SKA was shown to be involved in phosphorylating sphingosine whereas the tests did not exemplify phosphorylation by the cloned SKB and SKC genes.

Example 5 Antisense analysis

Knowledge of the correct, complete cDNA sequence of human SK enables its use as a tool for antisense technology in the investigation of gene function. Oligonucleotides, cDNA or genomic fragments comprising the antisense strand of SK are used either in vitro or in vivo to inhibit expression of the mRNA. Such technology is now well known in the art, and antisense molecules can be designed at various locations along the nucleotide sequences. By treatment of cells or whole test animals with such antisense sequences, the gene of interest is effectively turned off. Frequently, the function of the gene is ascertained by observing behavior at the intracellular, cellular, tissue or organismal level (e.g., lethality, loss of differentiated function, changes in morphology, etc.).

In addition to using sequences constructed to interrupt transcription of a particular open reading frame, modifications of gene expression is obtained by designing antisense sequences to intron regions, promoter/enhancer elements, or even to trans-acting regulatory genes. Similarly, inhibition is achieved using Hogeboom base-pairing methodology, also known as "triple helix" base pairing.

Example 6 Expression of Human SK

Expression of human SK is accomplished by subcloning the cDNAs into appropriate expression vectors and transfecting the vectors into analogous expression hosts for example E.Coli. In a particular case, the vector is engineered such that it contains a promoter for β-galactosidase, upstream of the cloning site, followed by

sequence containing the amino-terminal Met and the subsequent 7 residues of β-galactosidase. Immediately following these eight residues is an engineered bacteriophage promoter useful for artificial priming and transcription and for providing a number of unique endonuclease restriction sites for cloning.

Induction of the isolated, transfected bacterial strain with IPTG using standard methods produces a fusion protein corresponding to the first seven residues of β-galactosidase, about 15 residues of "linker", and the peptide encoded within the cDNA. Since cDNA clone inserts are generated by an essentially random process, there is one chance in three that the included cDNA will lie in the correct frame for proper translation. If the cDNA is not in the proper reading frame, it is obtained by deletion or insertion of the appropriate number of bases using well known methods including in vitro mutagenesis, digestion with exonuclease III or mung bean nuclease, or the inclusion of an oligonucleotide linker of appropriate length.

The human SK cDNA is shuttled into other vectors known to be useful for expression of protein in specific hosts. Oligonucleotide primers containing cloning sites as well as a segment of DNA (about 25 bases) sufficient to hybridize to stretches at both ends of the target cDNA is synthesized chemically by standard methods. These primers are then used to amplify the desired gene segment by PCR. The resulting gene segment is digested with appropriate restriction enzymes under standard conditions and isolated by gel electrophoresis. Alternately, similar gene segments are produced by digestion of the cDNA with appropriate restriction enzymes. Using appropriate primers, segments of coding sequence from more than one gene are ligated together and cloned in appropriate vectors. It is possible to optimize expression by construction of such chimeric sequences.

Suitable expression hosts for such chimeric molecules include, but are not limited to, mammalian cells such as Chinese Hamster Ovary (CHO) and human 293 cells, insect cells such as Sf9 cells, yeast cells such as Saccharomyces cerevisiae, and bacteria such as E. coli. For each of these cell systems, a useful expression vector

also includes an origin of replication to allow propagation in bacteria and a selectable marker such as the β-lactamase antibiotic resistance gene to allow plasmid selection in bacteria. In addition, the vector may include a second selectable marker such as the neomycin phosphotransferase gene to allow selection in transfected eukaryotic host cells. Vectors for use in eukaryotic expression hosts require RNA processing elements such as 3' polyadenylation sequences if such are not part of the cDNA of interest.

Additionally, the vector contains promoters or enhancers which increase gene expression. Such promoters are host specific and include MMTV, SV40, and metallothionine promoters for CHO cells; trp, lac, tac and T7 promoters for bacterial hosts; and alpha factor, alcohol oxidase and PGH promoters for yeast. Transcription enhancers, such as the rous sarcoma virus enhancer, are used in mammalian host cells. Once homogeneous cultures of recombinant cells are obtained through standard culture methods, large quantities of recombinantly produced human SK are recovered from the conditioned medium and analyzed using chromatographic methods known in the art. For example, human SK can be expressibly cloned into the expression vector pcDNA3. This product can be used to transform, for example, HEK293 or COS by methodology standard in the art. Specifically, for example, using Lipofectamine (Gibco BRL catalog no. 18324-020) mediated gene transfer.

Example 7 Isolation of Recombinant SK

Human SK is expressed as a chimeric protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle WA). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen) between the purification domain and the human SK

sequence is useful to facilitate expression of human SK.

Example 8 Production of SK Specific Antibodies

Two approaches are utilized to raise antibodies to human SK, and each approach is useful for generating either polyclonal or monoclonal antibodies. In one approach, denatured protein from reverse phase HPLC separation is obtained in quantities up to 75 mg. This denatured protein is used to immunize mice or rabbits using standard protocols; about 100 micrograms are adequate for immunization of a mouse, while up to 1 mg might be used to immunize a rabbit. For identifying mouse hybridomas, the denatured protein is radioiodinated and used to screen potential murine B-cell hybridomas for those which produce antibody. This procedure requires only small quantities of protein, such that 20 mg is sufficient for labeling and screening of several thousand clones.

In the second approach, the amino acid sequence of an appropriate human SK domain, as deduced from translation of the cDNA, is analyzed to determine regions of high antigenicity. Oligopeptides comprising appropriate hydrophilic regions are synthesized and used in suitable immunization protocols to raise antibodies. Analysis to select appropriate epitopes is described by Ausubel FM et al (supra). The optimal amino acid sequences for immunization are usually at the C-terminus, the N-terminus and those intervening, hydrophilic regions of the polypeptide which are likely to be exposed to the external environment when the protein is in its natural conformation.

Typically, selected peptides, about 15 residues in length, are synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry and coupled to keyhole limpet hemocyanin (KLH; Sigma, St. Louis MO) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel FM et al, supra). If necessary, a cysteine is introduced at the N-terminus of the peptide to permit coupling to KLH. Rabbits are immunized with the peptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity by binding

the peptide to plastic, blocking with 1% bovine sewm albumin, reacting with antisera, washing and reacting with labeled (radioactive or fluorescent), affinity purified, specific goat anti-rabbit lgG.

Hybridomas are prepared and screened using standard techniques. Hybridomas of interest are detected by screening with labeled human SK to identify those fusions producing the monoclonal antibody with the desired specificity. In a typical protocol, wells of plates (FAST; Becton-Dickinson, Palo Alto CA) are coated during incubation with affinity purified, specific rabbit anti-mouse (or suitable antispecies lg) antibodies at 10 mg/ml. The coated wells are blocked with 1% BSA, washed and incubated with supernatants from hybridomas. After washing the wells are incubated with labeled human SK at 1 mg/ml. Supernatants with specific antibodies bind more labeled human SK than is detectable in the background. Then clones producing specific antibodies are expanded and subjected to two cycles of cloning at limiting dilution. Cloned hybridomas are injected into pristane-treated mice to produce ascites, and monoclonal antibody is purified from mouse ascetic fluid by affinity chromatography on Protein A. Monoclonal antibodies with affinities of at least 10^8 M^{-1}, preferably 10^9 to 10^{10} or stronger, are typically made by standard procedures as described in Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; and in Goding (1986) Monoclonal Antibodies: Principles and Practice, Academic Press, New York City, both incorporated herein by reference.

Example 9 Diagnostic Test Using Human SK Specific Antibodies

Particular Human SK antibodies are useful for investigating signal transduction and the diagnosis of infectious or hereditary conditions which are characterized by differences in the amount or distribution of human SK or downstream products of an active signaling cascade.

Diagnostic tests for human SK include methods utilizing antibody and a label to detect human SK in human body fluids, membranes, cells, tissues or extracts of such. The polypeptides and antibodies of the present invention are used with or without modification. Frequently, the polypeptides and antibodies are labeled by joining them, either covalently or noncovalently, with a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, chromogenic agents, magnetic particles and the like. Patents teaching the use of such labels include US Patent No's. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in US Patent No.4,816,567, Incorporated herein by reference.

A variety of protocols for measuring soluble or membrane-bound human SK, using either polyclonal or monoclonal antibodies specific for the protein, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on human SK is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, DE et al (1983, J Exp. Med. 158:1211f).

Example 10 Purification of Native Human SK Using Specific Antibodies

Native or recombinant human SK is purified by immunoaffinity chromatography using antibodies specific for human SK. In general, an immunoaffinity column is constructed by covalently coupling the anti-TRH antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway NJ). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated Sepharose (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such immunoaffinity columns are utilized in the purification of human SK by preparing a fraction from cells containing human SK in a soluble form. This preparation is derived by solubilization of whole cells or of a subcellular fraction obtained via differential centrifugation (with or without addition of detergent) or by other methods well known in the art. Alternatively, soluble human SK containing a signal sequence is secreted in useful quantity into the medium in which the cells are grown.

Example 11 Drug Screening

This invention is particularly useful for screening therapeutic compounds by using human SK or binding fragments thereof in any of a variety of drug screening techniques. For example, human SK activity can be measured using any of a variety of appropriate functional assays in which activation of the kinase results in an observable change in the level of a particular produc. Thus, the present invention provides methods of screening for drugs or any other agents which are affect by humand SK.

Alternatively, the polypeptide or fragment employed in such a test is either free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells (or membrane preparations therefrom) which are stably transformed with

recombinant nucleic acids expressing the polypeptide or fragment. Drugs are screened against such transformed cells. One measures, for example, the formation of a phosphorylated product and compares that with a control.

Example 12 Rational Drug Design

Herein, the goal of rational drug design is to produce structural analogs of biologically active lipids of interest or of small molecules with which they interact, agonists, antagonists, or inhibitors.

In one approach, the three-dimensional structure of a protein of interest, or of a protein-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of a polypeptide is gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design efficient inhibitors. Useful examples of rational drug design includes molecules which have improved activity or stability as shown by Braxton S and Wells JA (1992, Biochemistry 31:7796-7801) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda SB et al (1993 J Biochem 113:742-46), incorporated herein by reference.

Example 13 Use and Administration of Antibodies, Inhibitors, or Antagonists

Antibodies, inhibitors, or antagonists of human SK (or other treatments to limit signal transduction, LST) provide different effects when administered therapeutically. LSTs are formulated in a nontoxic, inert, pharmaceutically acceptable aqueous carrier medium preferably at a pH of about 5 to 8, more preferably 6 to 8, although pH may vary according to the characteristics of the

antibody, inhibitor, or antagonist being formulated and the condition to be treated. Characteristics of LSTs include solubility of the molecule, half-life and antigenicity/immunogenicity. These and other characteristics aid in defining an effective carrier.

LSTs are delivered by known routes of administration including but not limited to topical creams and gels; transmucosal spray and aerosol; transdermal patch and bandage; injectable, intravenous and lavage formulations; and orally administered liquids and pills particularly formulated to resist stomach acid and enzymes. The particular formulation, exact dosage, and route of administration is determined by the attending physician and varies according to each specific situation.

Such determinations are made by considering multiple variables such as the condition to be treated, the LST to be administered, and the pharmacokinetic profile of a particular LST. Additional factors which are taken into account include severity of the disease state, patient's age, weight, gender and diet, time and frequency of LST administration, possible combination with other drugs, reaction sensitivities, and tolerance/response to therapy. Long acting LST formulations might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular LST.

Normal dosage amounts vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see US Patent Nos. 4,657,760; 5,206,344; or 5,225,212. Those skilled in the art employ different formulations for different LSTs. Administration to cells such as nerve cells necessitates delivery in a manner different from that to other cells such as vascular endothelial cells.

It is contemplated that abnormal signal transduction, trauma, or diseases which trigger humans SK activity are treatable with LSTs. These conditions or

diseases are specifically diagnosed by the tests discussed above, and such testing should be performed in suspected cases of viral, bacterial or fungal infections: allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or other conditions which activate the genes of lymphoid or neuronal tissues.

All publications and patent applications mentioned herein are incorporated by reference for the purpose of describing the methodologies, cell lines and vectors, among other things. However, nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure, for example, by virtue of prior invention.

Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments.

Claims

1. An isolated polynucledotide comprising a sequence encoding human sphingosine kinase as selected from the group consisting of:
   (a) human sphingosine kinase A or variants thereof;
   (b) human sphingosine kinase B or variants thereof; and
   (c) human sphingosine kinase C or variants thereof.

2. The isolated polynucleotide of claim 1 wherein said sequence encodes human sphingosine kinase A or variants thereof.

3. The isolated polynucleotide of claim 2 wherein said sequence encodes the sequence of Figure 3.

4. The isolated polynucleotide of claim 1 wherein said sequence encodes human sphingosine kinase B or variants thereof.

5. The isolated polynucleotide of claim 4 wherein said sequence encodes the sequence of Figure 6.

6. The isolated polynucleotide of claim 1 wherein said sequence encodes human sphingosine kinase C or variants thereof.

7. The isolated polynucleotide of claim 6 wherein said sequence encodes the sequence of Figure 9.

8. An isolated polynucleotide sequence comprising a complement of claim 1.

9. A composition containing the isolated polynucleotide sequence of claim 8 and an acceptable excipient.

10. An expression construct containing the isolated polynucleotide of claim 1.

11. A host cell containing the expression vector of claim 10.

12. A method for making a purified polypeptide comprising the amino acid sequence for human sphingosine kinase wherein the method comprises the steps of culturing a host cell of claim 11 in suitable conditions to express said polypeptide and isolating and purifying said expressed polypeptide.

13. A purified polypeptide comprising the amino acid sequence for human sphingosine kinase prepared by the method of claim 12.

14. A method of screening a compound for determining the capability of said compound to inhibit or activate human sphingosine kinase activity, which method comprises:

contacting a host cell of claim 11 with said compound; and

measuring the inhibition or activation of human sphingosine kinase activity.

Figure 1 Full-length human PSKA cDNA.

```
    GAGGTTATGGATCCAGCGGGCGGCCCCCGGGGCGTGCTCCCGCGGCCCTGCCGCGTGCTG
1   ---------+---------+---------+---------+---------+---------+   60
    CTCCAATACCTAGGTCGCCCGCCGGGGGCCCCGCACGAGGGCGCCGGGACGGCGCACGAC

    GTGCTGCTGAACCCGCGCGGCGGCAAGGGCAAGGCCTTGCAGCTCTTCCGGAGTCACGTG
61  ---------+---------+---------+---------+---------+---------+  120
    CACGACGACTTGGGCGCGCCGCCGTTCCCGTTCCGGAACGTCGAGAAGGCCTCAGTGCAC

    CAGCCCCTTTTGGCTGAGGCTGAAATCTCCTTCACGCTGATGCTCACTGAGCGGCGGAAC
121 ---------+---------+---------+---------+---------+---------+  180
    GTCGGGGAAAACCGACTCCGACTTTAGAGGAAGTGCGACTACGAGTGACTCGCCGCCTTG

    CACGCGCGGGAGCTGGTGCGGTCGGAGGAGCTGGGCCGCTGGGACGCTCTGGTGGTCATG
181 ---------+---------+---------+---------+---------+---------+  240
    GTGCGCGCCCTCGACCACGCCAGCCTCCTCGACCCGGCGACCCTGCGAGACCACCAGTAC

    TCTGGAGACGGGCTGATGCACGAGGTGGTGAACGGGCTCATGGAGCGGCCTGACTGGGAG
241 ---------+---------+---------+---------+---------+---------+  300
    AGACCTCTGCCCGACTACGTGCTCCACCACTTGCCCGAGTACCTCGCCGGACTGACCCTC

    ACCGCCATCCAGAAaCCCCTGTGTATCCTCCCAGCAGGCTCTGGCAACGCGCTGGCAGCT
301 ---------+---------+---------+---------+---------+---------+  360
    TGGCGGTAGGTCTTtGGGGACACATAGGAGGGTCGTCCGAGACCGTTGCGCGACCGTCGA

    TCCTTGAACCATTATGCTGGCTATGAGCAGGTCACCAATGAAGACCTCCTGACCAACTGC
361 ---------+---------+---------+---------+---------+---------+  420
    AGGAACTTGGTAATACGACCGATACTCGTCCAGTGGTTACTTCTGGAGGACTGGTTGACG

    ACGCTATTGCTGTGCCGCCGGCTGCTGTCACCCATGAACCTGCTGTCTCTGCACACGGCT
421 ---------+---------+---------+---------+---------+---------+  480
    TGCGATAACGACACGGCGGCCGACGACAGTGGGTACTTGGACGACAGAGACGTGTGCCGA

    TCGGGGCTGCGCCTCTTCTCTGTGCTCAGCCTGGCCTGGGGCTTCATTGCTGATGTGGAC
481 ---------+---------+---------+---------+---------+---------+  540
    AGCCCCGACGCGGAGAAGAGACACGAGTCGGACCGGACCCCGAAGTAACGACTACACCTG

    CTAGAGAGTGAGAAGTATCGGCGTCTGGGGGAGATGCGCTTCACTCTGGGCACCTTCCTG
541 ---------+---------+---------+---------+---------+---------+  600
    GATCTCTCACTCTTCATAGCCGCAGACCCCCTCTACGCGAAGTGAGACCCGTGGAAGGAC

    CGTCTGGCAGCCCTGCGCACCTACCGCGGCCGACTGGCCTACCTCCCTGTAGGAAGAGTG
601 ---------+---------+---------+---------+---------+---------+  660
    GCAGACCGTCGGGACGCGTGGATGGCGCCGGCTGACCGGATGGAGGGACATCCTTCTCAC

    GGTTCCAAGACACCTGCCTCCCCCGTTGTGGTCCAGCAGGGCCCGGTAGATGCACACCTT
661 ---------+---------+---------+---------+---------+---------+  720
    CCAAGGTTCTGTGGACGGAGGGGGCAACACCAGGTCGTCCCGGGCCATCTACGTGTGGAA

    GTGCCACTGGAGGAGCCAGTGCCCTCTCACTGGACAGTGGTGCCCGACGAGGACTTTGTG
721 ---------+---------+---------+---------+---------+---------+  780
    CACGGTGACCTCCTCGGTCACGGGAGAGTGACCTGTCACCACGGGCTGCTCCTGAAACAC
```

Figure 1 cont'd

```
      CTAGTCCTGGCACTGCTGCACTCGCACCTGGGCAGTGAGATGTTTGCTGCACCCATGGGC
 781  ---------+---------+---------+---------+---------+---------+  840
      GATCAGGACCGTGACGACGTGAGCGTGGACCCGTCACTCTACAAACGACGTGGGTACCCG

      CGCTGTGCAGCTGGCGTCATGCATCTGTTCTACGTGCGGGCGGGAGTGTCTCGTGCCATG
 841  ---------+---------+---------+---------+---------+---------+  900
      GCGACACGTCGACCGCAGTACGTAGACAAGATGCACGCCCGCCCTCACAGAGCACGGTAC

      CTGCTGCGCCTCTTCCTGGCCATGGAGAAGGGCAGGCATATGGAGTATGAATGCCCCTAC
 901  ---------+---------+---------+---------+---------+---------+  960
      GACGACGCGGAGAAGGACCGGTACCTCTTCCCGTCCGTATACCTCATACTTACGGGGATG

      TTGGTATATGTGCCCGTGGTCGCCTTCCGCTTGGAGCCCAAGGATGGGAAAGGTGTGTTT
 961  ---------+---------+---------+---------+---------+---------+ 1020
      AACCATATACACGGGCACCAGCGGAAGGCGAACCTCGGGTTCCTACCCTTTCCACACAAA

      GCAGTGGATGGGGAATTGATGGTTAGgCGAGCCGTGCAGGGCCAGGTGCACCCAAACTAC
1021  ---------+---------+---------+---------+---------+---------+ 1080
      CGTCACCTACCCCTTAACTACCAATCcGCTCGGCACGTCCCGGTCCACGTGGGTTTGATG

      TTCTGGATGGTCAGCGGTTGCGTGGAGCCCCCGCCCAGCTGGAAGCCCCAGCAGATGCCA
1081  ---------+---------+---------+---------+---------+---------+ 1140
      AAGACCTACCAGTCGCCAACGCACCTCGGGGGCGGGTCGACCTTCGGGGTCGTCTACGGT

      CCGCCAGAAGAGCCCTTATGACCCCTGGGCCGCGCTGTGCCTTAGTGTCTACTTGCAGGA
1141  ---------+---------+---------+---------+---------+---------+ 1200
      GGCGGTCTTCTCGGGAATACTGGGGACCCGGCGCGACACGGAATCACAGATGAACGTCCT

      CCCTTCCTCCTTCCCTAGGGCTGCAGGGCCTGTCCACAGCTCCTGTGGGGGTGGAGGAGA
1201  ---------+---------+---------+---------+---------+---------+ 1260
      GGGAAGGAGGAAGGGATCCCGACGTCCCGGACAGGTGTCGAGGACACCCCCACCTCCTCT

      CTCCTCTGGAGAAGGGTGAGAAGGTGGAGGCTATGCTTTGGGGGGACAGGCCAGAATGAA
1261  ---------+---------+---------+---------+---------+---------+ 1320
      GAGGAGACCTCTTCCCACTCTTCCACCTCCGATACGAAACCCCCCTGTCCGGTCTTACTT

      GTCCTGGGTCAGGAGCCCAGCTGGCTGGGCCCAGCTGCCTATGTAAGGCCTTCTAGTTTG
1321  ---------+---------+---------+---------+---------+---------+ 1380
      CAGGACCCAGTCCTCGGGTCGACCGACCCGGGTCGACGGATACATTCCGGAAGATCAAAC

      TTCTGAGACCCCCACCCCACGAACCAAATCCAAATAAAGTGACATTCCCAAAAAAAAAAA
1381  ---------+---------+---------+---------+---------+---------+ 1440
      AAGACTCTGGGGGTGGGGTGCTTGGTTTAGGTTTATTTCACTGTAAGGGTTTTTTTTTTT

      AAAAAAA
1441  -------  1447
      TTTTTTT
```

Figure 2. Coding region of human PSKA cDNA sequence.

```
    ATGGATCCAGCGGGCGGCCCCCGGGGCGTGCTCCCGCGGCCCTGCCGCGTGCTGGTGCTG
  7 ---+---------+---------+---------+---------+---------+---------+------   66
    TACCTAGGTCGCCCGCCGGGGGCCCCGCACGAGGGCGCCGGGACGGCGCACGACCACGAC

    CTGAACCCGCGCGGCGGCAAGGGCAAGGCCTTGCAGCTCTTCCGGAGTCACGTGCAGCCC
 67 ---+---------+---------+---------+---------+---------+---------+------  126
    GACTTGGGCGCGCCGCCGTTCCCGTTCCGGAACGTCGAGAAGGCCTCAGTGCACGTCGGG

    CTTTTGGCTGAGGCTGAAATCTCCTTCACGCTGATGCTCACTGAGCGGCGGAACCACGCG
127 ---+---------+---------+---------+---------+---------+---------+------  186
    GAAAACCGACTCCGACTTTAGAGGAAGTGCGACTACGAGTGACTCGCCGCCTTGGTGCGC

    CGGGAGCTGGTGCGGTCGGAGGAGCTGGGCCGCTGGGACGCTCTGGTGGTCATGTCTGGA
187 ---+---------+---------+---------+---------+---------+---------+------  246
    GCCCTCGACCACGCCAGCCTCCTCGACCCGGCGACCCTGCGAGACCACCAGTACAGACCT

    GACGGGCTGATGCACGAGGTGGTGAACGGGCTCATGGAGCGGCCTGACTGGGAGACCGCC
247 ---+---------+---------+---------+---------+---------+---------+------  306
    CTGCCCGACTACGTGCTCCACCACTTGCCCGAGTACCTCGCCGGACTGACCCTCTGGCGG

    ATCCAGAAaCCCCTGTGTATCCTCCCAGCAGGCTCTGGCAACGCGCTGGCAGCTTCCTTG
307 ---+---------+---------+---------+---------+---------+---------+------  366
    TAGGTCTTtGGGGACACATAGGAGGGTCGTCCGAGACCGTTGCGCGACCGTCGAAGGAAC

    AACCATTATGCTGGCTATGAGCAGGTCACCAATGAAGACCTCCTGACCAACTGCACGCTA
367 ---+---------+---------+---------+---------+---------+---------+------  426
    TTGGTAATACGACCGATACTCGTCCAGTGGTTACTTCTGGAGGACTGGTTGACGTGCGAT

    TTGCTGTGCCGCCGGCTGCTGTCACCCATGAACCTGCTGTCTCTGCACACGGCTTCGGGG
427 ---+---------+---------+---------+---------+---------+---------+------  486
    AACGACACGGCGGCCGACGACAGTGGGTACTTGGACGACAGAGACGTGTGCCGAAGCCCC

    CTGCGCCTCTTCTCTGTGCTCAGCCTGGCCTGGGGCTTCATTGCTGATGTGGACCTAGAG
487 ---+---------+---------+---------+---------+---------+---------+------  546
    GACGCGGAGAAGAGACACGAGTCGGACCGGACCCCGAAGTAACGACTACACCTGGATCTC

    AGTGAGAAGTATCGGCGTCTGGGGGAGATGCGCTTCACTCTGGGCACCTTCCTGCGTCTG
547 ---+---------+---------+---------+---------+---------+---------+------  606
    TCACTCTTCATAGCCGCAGACCCCCTCTACGCGAAGTGAGACCCGTGGAAGGACGCAGAC

    GCAGCCCTGCGCACCTACCGCGGCCGACTGGCCTACCTCCCTGTAGGAAGAGTGGGTTCC
607 ---+---------+---------+---------+---------+---------+---------+------  666
    CGTCGGGACGCGTGGATGGCGCCGGCTGACCGGATGGAGGGACATCCTTCTCACCCAAGG

    AAGACACCTGCCTCCCCCGTTGTGGTCCAGCAGGGCCCGGTAGATGCACACCTTGTGCCA
667 ---+---------+---------+---------+---------+---------+---------+------  726
    TTCTGTGGACGGAGGGGGCAACACCAGGTCGTCCCGGGCCATCTACGTGTGGAACACGGT

    CTGGAGGAGCCAGTGCCCTCTCACTGGACAGTGGTGCCCGACGAGGACTTTGTGCTAGTC
727 ---+---------+---------+---------+---------+---------+---------+------  786
    GACCTCCTCGGTCACGGGAGAGTGACCTGTCACCACGGGCTGCTCCTGAAACACGATCAG

    CTGGCACTGCTGCACTCGCACCTGGGCAGTGAGATGTTTGCTGCACCCATGGGCCGCTGT
787 ---+---------+---------+---------+---------+---------+---------+------  846
    GACCGTGACGACGTGAGCGTGGACCCGTCACTCTACAAACGACGTGGGTACCCGGCGACA

    GCAGCTGGCGTCATGCATCTGTTCTACGTGCGGGCGGGAGTGTCTCGTGCCATGCTGCTG
```

Figure 2 cont'd

```
847  ---+---------+---------+---------+---------+---------+---------+------  906
     CGTCGACCGCAGTACGTAGACAAGATGCACGCCCGCCCTCACAGAGCACGGTACGACGAC

     CGCCTCTTCCTGGCCATGGAGAAGGGCAGGCATATGGAGTATGAATGCCCCTACTTGGTA
907  ---+---------+---------+---------+---------+---------+---------+------  966
     GCGGAGAAGGACCGGTACCTCTTCCCGTCCGTATACCTCATACTTACGGGGATGAACCAT

     TATGTGCCCGTGGTCGCCTTCCGCTTGGAGCCCAAGGATGGGAAAGGTGTGTTTGCAGTG
967  ---+---------+---------+---------+---------+---------+---------+------  1026
     ATACACGGGCACCAGCGGAAGGCGAACCTCGGGTTCCTACCCTTTCCACACAAACGTCAC

     GATGGGGAATTGATGGTTAGgCGAGCCGTGCAGGGCCAGGTGCACCCAAACTACTTCTGG
1027 ---+---------+---------+---------+---------+---------+---------+------  1086
     CTACCCCTTAACTACCAATCcGCTCGGCACGTCCCGGTCCACGTGGGTTTGATGAAGACC

     ATGGTCAGCGGTTGCGTGGAGCCCCCGCCCAGCTGGAAGCCCCAGCAGATGCCACCGCCA
1087 ---+---------+---------+---------+---------+---------+---------+------  1146
     TACCAGTCGCCAACGCACCTCGGGGGCGGGTCGACCTTCGGGGTCGTCTACGGTGGCGGT

     GAAGAGCCCTTATGA
1147 ---+---------+-  1161
     CTTCTCGGGAATACT
```

Figure 3. Predicted amino acid sequence of human PSKA protein.

```
    MDPAGGPRGVLPRPCRVLVLLNPRGGKGKALQLFRSHVQPLLAEAEISFTLMLTERRNHA
 1  ---------+---------+---------+---------+---------+---------+   60

    RELVRSEELGRWDALVVMSGDGLMHEVVNGLMERPDWETAIQKPLCILPAGSGNALAASL
61  ---------+---------+---------+---------+---------+---------+  120

    NHYAGYEQVTNEDLLTNCTLLLCRRLLSPMNLLSLHTASGLRLFSVLSLAWGFIADVDLE
121 ---------+---------+---------+---------+---------+---------+  180

    SEKYRRLGEMRFTLGTFLRLAALRTYRGRLAYLPVGRVGSKTPASPVVVQQGPVDAHLVP
181 ---------+---------+---------+---------+---------+---------+  240

    LEEPVPSHWTVVPDEDFVLVLALLHSHLGSEMFAAPMGRCAAGVMHLFYVRAGVSRAMLL
241 ---------+---------+---------+---------+---------+---------+  300

    RLFLAMEKGRHMEYECPYLVYVPVVAFRLEPKDGKGVFAVDGELMVRRAVQGQVHPNYFW
301 ---------+---------+---------+---------+---------+---------+  360

    MVSGCVEPPPSWKPQQMPPPEEPL*
361 ---------+---------+-----  385
```

Figure 4. Sequence of full-length cDNA encoding human PSKB.

```
    AGCCGCGAGCTGGACCAGCCGTGCAAATCTCTAGAAGATGACGGTGTTCTTTAAAACGCT
1   ---------+---------+---------+---------+---------+---------+   60
    TCGGCGCTCGACCTGGTCGGCACGTTTAGAGATCTTCTACTGCCACAAGAAATTTTGCGA

    TCGAAATCACTGGAAGAAAACTACAGCTGGGCTCTGCCTGCTGACCTGGGGAGGCCATTG
61  ---------+---------+---------+---------+---------+---------+  120
    AGCTTTAGTGACCTTCTTTTGATGTCGACCCGAGACGGACGACTGGACCCCTCCGGTAAC

    GCTCTATGGAAAACACTGTGATAACCTCCTAAGGAGAGCAGCCTGTCAAGAAGCTCAGGT
121 ---------+---------+---------+---------+---------+---------+  180
    CGAGATACCTTTTGTGACACTATTGGAGGATTCCTCTCGTCGGACAGTTCTTCGAGTCCA

    GTTTGGCAATCAACTCATTCCTCCCAATGCACAAGTGAAGAAGGCCACTGTTTTTCTCAA
181 ---------+---------+---------+---------+---------+---------+  240
    CAAACCGTTAGTTGAGTAAGGAGGGTTACGTGTTCACTTCTTCCGGTGACAAAAAGAGTT

    TCCTGCAGCTTGCAAAGGAAAAGCCAGGACTCTATTTGAAAAAAATGCTGCCCCGATTTT
241 ---------+---------+---------+---------+---------+---------+  300
    AGGACGTCGAACGTTTCCTTTTCGGTCCTGAGATAAACTTTTTTTACGACGGGGCTAAAA

    ACATTTATCTGGCATGGATGTGACTATTGTTAAGACAGATTATGAGGGACAAGCCAAGAA
301 ---------+---------+---------+---------+---------+---------+  360
    TGTAAATAGACCGTACCTACACTGATAACAATTCTGTCTAATACTCCCTGTTCGGTTCTT

    ACTCCTGGAACTGATGGAAAACACGGATGTGATCATTGTTGCAGGAGGAGATGGGACACT
361 ---------+---------+---------+---------+---------+---------+  420
    TGAGGACCTTGACTACCTTTTGTGCCTACACTAGTAACAACGTCCTCCTCTACCCTGTGA

    GCAGGAGGTTGTTACTGGTGTTCTTCGACGAACAGATGAGGCTACCTTCAGTAAGATTCC
421 ---------+---------+---------+---------+---------+---------+  480
    CGTCCTCCAACAATGACCACAAGAAGCTGCTTGTCTACTCCGATGGAAGTCATTCTAAGG

    CATTGGATTTATCCCACTGGGAGAGACCAGTAGTTTGAGTCATACCCTCTTTGCCGAAAG
481 ---------+---------+---------+---------+---------+---------+  540
    GTAACCTAAATAGGGTGACCCTCTCTGGTCATCAAACTCAGTATGGGAGAAACGGCTTTC

    TGGAAACAAAGTCCAACATATTACTGATGCCACACTTGCCATTGTGAAAGGAGAGACAGT
541 ---------+---------+---------+---------+---------+---------+  600
    ACCTTTGTTTCAGGTTGTATAATGACTACGGTGTGAACGGTAACACTTTCCTCTCTGTCA

    TCCACTTGATTTCTTGCAGATCAAGGGTGAAAAGGAACAGCCTGTATTTGCAATGACCGG
601 ---------+---------+---------+---------+---------+---------+  660
    AGGTGAACTAAAGAACGTCTAGTTCCCACTTTTCCTTGTCGGACATAAACGTTACTGGCC

    CCTTCGATGGGGATCTTTCAGAGATGCTGGCGTCAAAGTTAGCAAGTACTGGTATCTTGG
661 ---------+---------+---------+---------+---------+---------+  720
    GGAAGCTACCCCTAGAAAGTCTCTACGACCGCAGTTTCAATCGTTCATGACCATAGAACC

    GCCTCTAAAAATCAAAGCAGCCCACTTTTTCAGCACTCTTAAGGAGTGGCCTCAGACTCA
721 ---------+---------+---------+---------+---------+---------+  780
    CGGAGATTTTTAGTTTCGTCGGGTGAAAAAGTCGTGAGAATTCCTCACCGGAGTCTGAGT

    TCAAGCCTCTATCTCATACACGGGACCTACAGAGAGACCTCCCAATGAACCAGAGGAGAC
781 ---------+---------+---------+---------+---------+---------+  840
    AGTTCGGAGATAGAGTATGTGCCCTGGATGTCTCTCTGGAGGGTTACTTGGTCTCCTCTG

    CCCTGTACAAAGGCCTTCTTTGTACAGGAGAATATTACGAAGGCTTGCGTCCTACTGGGC
```

Figure 4 cont'd


```
841  ---------+---------+---------+---------+---------+---------+  900
     GGGACATGTTTCCGGAAGAAACATGTCCTCTTATAATGCTTCCGAACGCAGGATGACCCG

     ACAACCACAGGATGCCCTTTCCCAAGAGGTGAGCCCGGAGGTCTGGAAAGATGTGCAGCT
901  ---------+---------+---------+---------+---------+---------+  960
     TGTTGGTGTCCTACGGGAAAGGGTTCTCCACTCGGGCCTCCAGACCTTTCTACACGTCGA

     GTCCACCATTGAACTGTCCATCACAACACGGAATAATCAGCTTGACCCGACAAGCAAAGA
961  ---------+---------+---------+---------+---------+---------+ 1020
     CAGGTGGTAACTTGACAGGTAGTGTTGTGCCTTATTAGTCGAACTGGGCTGTTCGTTTCT

     AGATTTTCTGAATATCTGCATTGAACCTGACACCATCAGCAAAGGAGACTTTATAACTAT
1021 ---------+---------+---------+---------+---------+---------+ 1080
     TCTAAAAGACTTATAGACGTAACTTGGACTGTGGTAGTCGTTTCCTCTGAAATATTGATA

     AGGAAGTCGAAAGGTGAGAAACCCCAAGCTGCACGTGGAGGGCACGGAGTGTCTCCAAGC
1081 ---------+---------+---------+---------+---------+---------+ 1140
     TCCTTCAGCTTTCCACTCTTTGGGGTTCGACGTGCACCTCCCGTGCCTCACAGAGGTTCG

     CAGCCAGTGCACTTTGCTTATCCCGGAGGGAGCAGGGGGCTCTTTTAGCATTGACAGTGA
1141 ---------+---------+---------+---------+---------+---------+ 1200
     GTCGGTCACGTGAAACGAATAGGGCCTCCCTCGTCCCCCGAGAAAATCGTAACTGTCACT

     GGAGTATGAAGCGATGCCTGTGGAGGTGAAACTGCTCCCCAGGAAGCTGCAGTTCTTCTG
1201 ---------+---------+---------+---------+---------+---------+ 1260
     CCTCATACTTCGCTACGGACACCTCCACTTTGACGAGGGGTCCTTCGACGTCAAGAAGAC

     TGATCCTAGGAAGAGAGAACAGATGCTCACAAGCCCCACCCAGTGAGCAGCAGAAGACAA
1261 ---------+---------+---------+---------+---------+---------+ 1320
     ACTAGGATCCTTCTCTCTTGTCTACGAGTGTTCGGGGTGGGTCACTCGTCGTCTTCTGTT

     GCACTCTGAGACCACACTTTAGGCCACCGGTGGGACCAAAAGGGAACAGGTGCCTCAGCC
1321 ---------+---------+---------+---------+---------+---------+ 1380
     CGTGAGACTCTGGTGTGAAATCCGGTGGCCACCCTGGTTTTCCCTTGTCCACGGAGTCGG

     ATCCCAACAGTGTCGTCAGAGGGTCCCCAGGGCATTTTCATGGCAAGTACCCCTCTGCCC
1381 ---------+---------+---------+---------+---------+---------+ 1440
     TAGGGTTGTCACAGCAGTCTCCCAGGGGTCCCGTAAAAGTACCGTTCATGGGGAGACGGG

     CCACTCCAGCAGTGCTTCCCAAAGTGTGCTCTGTCACCTGCTTTGCAATCGGCTTCCATT
1441 ---------+---------+---------+---------+---------+---------+ 1500
     GGTGAGGTCGTCACGAAGGGTTTCACACGAGACAGTGGACGAAACGTTAGCCGAAGGTAA

     AGCGCATGTTTTATTTTGGTGTGACGGTTGGCCCTCCTAAACACGGACTTTCCTCAGGCT
1501 ---------+---------+---------+---------+---------+---------+ 1560
     TCGCGTACAAAATAAAACCACACTGCCAACCGGGAGGATTTGTGCCTGAAAGGAGTCCGA

     GGTTCAAGACGGAAAAGGACTTTCTTCTGTTTTCTTCCAAAGTGCAACCACAGTGGAGAG
1561 ---------+---------+---------+---------+---------+---------+ 1620
     CCAAGTTCTGCCTTTTCCTGAAAGAAGACAAAAGAAGGTTTCACGTTGGTGTCACCTCTC

     CCCACGGTGGGCTTAGCCTGCCTAGGCCCTTCCATTTCTCTTCTTTGACCGTGCTAGGAA
1621 ---------+---------+---------+---------+---------+---------+ 1680
     GGGTGCCACCCGAATCGGACGGATCCGGGAAGGTAAAGAGAAGAAACTGGCACGATCCTT

     TTCCAGGAAAGTGCATTCCTGCCCTGGTGACCTTTTCCTATGTCTAGGCTCCTCCACAGG  -
1681 ---------+---------+---------+---------+---------+---------+ 1740
     AAGGTCCTTTCACGTAAGGACGGGACCACTGGAAAAGGATACAGATCCGAGGAGGTGTCC

     TGCTGCTATTTTGTGAGCTCCGGCTCCTGTTTAGCTTTTATTTCAGTTCTAACCTCAGTC
```

Figure 4 cont'd


```
1741  ---------+---------+---------+---------+---------+---------+  1800
      ACGACGATAAAACACTCGAGGCCGAGGACAAATCGAAAATAAAGTCAAGATTGGAGTCAG

      CAGAAACATATGTGAGGTTGTTTCCCTCTTCAGCCACGGCTACAATACCGGAAAATGCTA
1801  ---------+---------+---------+---------+---------+---------+  1860
      GTCTTTGTATACACTCCAACAAAGGGAGAAGTCGGTGCCGATGTTATGGCCTTTTACGAT

      GTTTTTATTTATTTTTTTAAGTAGTGCTTCCTAAATGGTTTGCATGAGAGCCACCTGGGG
1861  ---------+---------+---------+---------+---------+---------+  1920
      CAAAAATAAATAAAAAAATTCATCACGAAGGATTTACCAAACGTACTCTCGGTGGACCCC

      TACATGTTGAAAACTTATTTGGGGTCTACCCCAAACCTAATAACCCAAATTTGGGGATGG
1921  ---------+---------+---------+---------+---------+---------+  1980
      ATGTACAACTTTTGAATAAACCCCAGATGGGGTTTGGATTATTGGGTTTAAACCCCTACC

      GGCCCAGGAATATGCATTTTTAAAAAGTCATCTGCCCTTCCCAGGTGATTCTGTAAGTTG
1981  ---------+---------+---------+---------+---------+---------+  2040
      CCGGGTCCTTATACGTAAAAATTTTTCAGTAGACGGGAAGGGTCCACTAAGACATTCAAC

      TCCCTCAACTGTACTTGGAGAAATCGTGTTTTAAAGCAGTAGTCCACAAAGTATTCTGCT
2041  ---------+---------+---------+---------+---------+---------+  2100
      AGGGAGTTGACATGAACCTCTTTAGCACAAAATTTCGTCATCAGGTGTTTCATAAGACGA

      CATGTGCCCCCAAAAGTATTTTGAAAAATCATGTATACCCTCACCCATCTAAGTTGATAT
2101  ---------+---------+---------+---------+---------+---------+  2160
      GTACACGGGGGTTTTCATAAAACTTTTTAGTACATATGGGAGTGGGTAGATTCAACTATA

      CTAAAATTTTATCTAAGTTGGTATCTAAAATTTTTCATGGGAAGTTAAATAGTTGACAAA
2161  ---------+---------+---------+---------+---------+---------+  2220
      GATTTTAAAATAGATTCAACCATAGATTTTAAAAAGTACCCTTCAATTTATCAACTGTTT

      GTATGTATTTGCTGGTGTCGTGTAAATATTGGTATTTTAAAATAAAAACTGTTACATCAC
2221  ---------+---------+---------+---------+---------+---------+  2280
      CATACATAAACGACCACAGCACATTTATAACCATAAAATTTTATTTTTGACAATGTAGTG

      TAAWWWAAAAAAAAAAAAAAAAAAA
2281  ---------+---------+------  2306
      ATTWWWTTTTTTTTTTTTTTTTTTT
```

Figure 5. Coding region of human PSKB cDNA sequence.

```
    ATGACGGTGTTCTTTAAAACGCTTCGAAATCACTGGAAGAAAACTACAGCTGGGCTCTGC
 38 --+---------+---------+---------+---------+---------+---------+-------  97
    TACTGCCACAAGAAATTTTGCGAAGCTTTAGTGACCTTCTTTTGATGTCGACCCGAGACG

    CTGCTGACCTGGGGAGGCCATTGGCTCTATGGAAAACACTGTGATAACCTCCTAAGGAGA
 98 --+---------+---------+---------+---------+---------+---------+-------  157
    GACGACTGGACCCCTCCGGTAACCGAGATACCTTTTGTGACACTATTGGAGGATTCCTCT

    GCAGCCTGTCAAGAAGCTCAGGTGTTTGGCAATCAACTCATTCCTCCCAATGCACAAGTG
158 --+---------+---------+---------+---------+---------+---------+-------  217
    CGTCGGACAGTTCTTCGAGTCCACAAACCGTTAGTTGAGTAAGGAGGGTTACGTGTTCAC

    AAGAAGGCCACTGTTTTTCTCAATCCTGCAGCTTGCAAAGGAAAAGCCAGGACTCTATTT
218 --+---------+---------+---------+---------+---------+---------+-------  277
    TTCTTCCGGTGACAAAAAGAGTTAGGACGTCGAACGTTTCCTTTTCGGTCCTGAGATAAA

    GAAAAAAATGCTGCCCCGATTTTACATTTATCTGGCATGGATGTGACTATTGTTAAGACA
278 --+---------+---------+---------+---------+---------+---------+-------  337
    CTTTTTTTACGACGGGGCTAAAATGTAAATAGACCGTACCTACACTGATAACAATTCTGT

    GATTATGAGGGACAAGCCAAGAAACTCCTGGAACTGATGGAAAACACGGATGTGATCATT
338 --+---------+---------+---------+---------+---------+---------+-------  397
    CTAATACTCCCTGTTCGGTTCTTTGAGGACCTTGACTACCTTTTGTGCCTACACTAGTAA

    GTTGCAGGAGGAGATGGGACACTGCAGGAGGTTGTTACTGGTGTTCTTCGACGAACAGAT
398 --+---------+---------+---------+---------+---------+---------+-------  457
    CAACGTCCTCCTCTACCCTGTGACGTCCTCCAACAATGACCACAAGAAGCTGCTTGTCTA

    GAGGCTACCTTCAGTAAGATTCCCATTGGATTTATCCCACTGGGAGAGACCAGTAGTTTG
458 --+---------+---------+---------+---------+---------+---------+-------  517
    CTCCGATGGAAGTCATTCTAAGGGTAACCTAAATAGGGTGACCCTCTCTGGTCATCAAAC

    AGTCATACCCTCTTTGCCGAAAGTGGAAACAAAGTCCAACATATTACTGATGCCACACTT
518 --+---------+---------+---------+---------+---------+---------+-------  577
    TCAGTATGGGAGAAACGGCTTTCACCTTTGTTTCAGGTTGTATAATGACTACGGTGTGAA

    GCCATTGTGAAAGGAGAGACAGTTCCACTTGATTTCTTGCAGATCAAGGGTGAAAAGGAA
578 --+---------+---------+---------+---------+---------+---------+-------  637
    CGGTAACACTTTCCTCTCTGTCAAGGTGAACTAAAGAACGTCTAGTTCCCACTTTTCCTT

    CAGCCTGTATTTGCAATGACCGGCCTTCGATGGGGATCTTTCAGAGATGCTGGCGTCAAA
638 --+---------+---------+---------+---------+---------+---------+-------  697
    GTCGGACATAAACGTTACTGGCCGGAAGCTACCCCTAGAAAGTCTCTACGACCGCAGTTT

    GTTAGCAAGTACTGGTATCTTGGGCCTCTAAAAATCAAAGCAGCCCACTTTTTCAGCACT
698 --+---------+---------+---------+---------+---------+---------+-------  757
    CAATCGTTCATGACCATAGAACCCGGAGATTTTTAGTTTCGTCGGGTGAAAAAGTCGTGA

    CTTAAGGAGTGGCCTCAGACTCATCAAGCCTCTATCTCATACACGGGACCTACAGAGAGA
758 --+---------+---------+---------+---------+---------+---------+-------  817
    GAATTCCTCACCGGAGTCTGAGTAGTTCGGAGATAGAGTATGTGCCCTGGATGTCTCTCT

    CCTCCCAATGAACCAGAGGAGACCCCTGTACAAAGGCCTTCTTTGTACAGGAGAATATTA
818 --+---------+---------+---------+---------+---------+---------+-------  877
    GGAGGGTTACTTGGTCTCCTCTGGGGACATGTTTCCGGAAGAAACATGTCCTCTTATAAT

    CGAAGGCTTGCGTCCTACTGGGCACAACCACAGGATGCCCTTTCCCAAGAGGTGAGCCCG
878 --+---------+---------+---------+---------+---------+---------+-------  937
```

Figure 5 cont'd


```
      GCTTCCGAACGCAGGATGACCCGTGTTGGTGTCCTACGGGAAAGGGTTCTCCACTCGGGC

      GAGGTCTGGAAAGATGTGCAGCTGTCCACCATTGAACTGTCCATCACAACACGGAATAAT
938   --+---------+---------+---------+---------+---------+---------+-------  997
      CTCCAGACCTTTCTACACGTCGACAGGTGGTAACTTGACAGGTAGTGTTGTGCCTTATTA

      CAGCTTGACCCGACAAGCAAAGAAGATTTTCTGAATATCTGCATTGAACCTGACACCATC
998   --+---------+---------+---------+---------+---------+---------+-------  1057
      GTCGAACTGGGCTGTTCGTTTCTTCTAAAAGACTTATAGACGTAACTTGGACTGTGGTAG

      AGCAAAGGAGACTTTATAACTATAGGAAGTCGAAAGGTGAGAAACCCCAAGCTGCACGTG
1058  --+---------+---------+---------+---------+---------+---------+-------  1117
      TCGTTTCCTCTGAAATATTGATATCCTTCAGCTTTCCACTCTTTGGGGTTCGACGTGCAC

      GAGGGCACGGAGTGTCTCCAAGCCAGCCAGTGCACTTTGCTTATCCCGGAGGGAGCAGGG
1118  --+---------+---------+---------+---------+---------+---------+-------  1177
      CTCCCGTGCCTCACAGAGGTTCGGTCGGTCACGTGAAACGAATAGGGCCTCCCTCGTCCC

      GGCTCTTTTAGCATTGACAGTGAGGAGTATGAAGCGATGCCTGTGGAGGTGAAACTGCTC
1178  --+---------+---------+---------+---------+---------+---------+-------  1237
      CCGAGAAAATCGTAACTGTCACTCCTCATACTTCGCTACGGACACCTCCACTTTGACGAG

      CCCAGGAAGCTGCAGTTCTTCTGTGATCCTAGGAAGAGAGAACAGATGCTCACAAGCCCC
1238  --+---------+---------+---------+---------+---------+---------+-------  1297
      GGGTCCTTCGACGTCAAGAAGACACTAGGATCCTTCTCTCTTGTCTACGAGTGTTCGGGG

      ACCCAGTGA
1298  --+-------  1306
      TGGGTCACT
```

Figure 6. Predicted polypeptide sequence of human PSKB protein.

```
   MTVFFKTLRNHWKKTTAGLCLLTWGGHWLYGKHCDNLLRRAACQEAQVFGNQLIPPNAQV
1  ---------+---------+---------+---------+---------+---------+   60

   KKATVFLNPAACKGKARTLFEKNAAPILHLSGMDVTIVKTDYEGQAKKLLELMENTDVII
61 ---------+---------+---------+---------+---------+---------+  120

   VAGGDGTLQEVVTGVLRRTDEATFSKIPIGFIPLGETSSLSHTLFAESGNKVQHITDATL
121--------+---------+---------+---------+---------+---------+  180

   AIVKGETVPLDFLQIKGEKEQPVFAMTGLRWGSFRDAGVKVSKYWYLGPLKIKAAHFFST
181 ---------+---------+---------+---------+---------+---------+ 240

   LKEWPQTHQASISYTGPTERPPNEPEETPVQRPSLYRRILRRLASYWAQPQDALSQEVSP
241 ---------+---------+---------+---------+---------+---------+ 300

   EVWKDVQLSTIELSITTRNNQLDPTSKEDFLNICIEPDTISKGDFITIGSRKVRNPKLHV
301 ---------+---------+---------+---------+---------+---------+ 360

   EGTECLQASQCTLLIPEGAGGSFSIDSEEYEAMPVEVKLLPRKLQFFCDPRKREQMLTSP
361 ---------+---------+---------+---------+---------+---------+ 420

   TQ*
421--- 423
```

Figure 7. Full-length human PSKC cDNA sequence.

```
    ATCTGAGATCATCGCCGTTGAGGAAACAGACGTTCACGGGAAACATCAAGGCAGTGGAAA
1   ---------+---------+---------+---------+---------+---------+   60
    TAGACTCTAGTAGCGGCAACTCCTTTGTCTGCAAGTGCCCTTTGTAGTTCCGTCACCTTT

    ATGGCAGAAAATGGAAAAGCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCACGACG
61  ---------+---------+---------+---------+---------+---------+  120
    TACCGTCTTTTACCTTTTCGGAATGCGAAAATGTCAAGTGACACATTTCTCTCGTGCTGC

    GCACCGCTGGAAGTGGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTT
121 ---------+---------+---------+---------+---------+---------+  180
    CGTGGCGACCTTCACCCGCGTCCACTGAAAGACCACAGGTCTCCTCGTCGACACAGTGAA

    GTGGCTGCAGACCCTGCGGGAGATGCTGGAGAAGCTGACGTCCAGACCAAAGCATTTACT
181 ---------+---------+---------+---------+---------+---------+  240
    CACCGACGTCTGGGACGCCCTCTACGACCTCTTCGACTGCAGGTCTGGTTTCGTAAATGA

    GGTATTTATCAACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATATATGAAAGAAAAGT
241 ---------+---------+---------+---------+---------+---------+  300
    CCATAAATAGTTGGGCAAACCTCCTTTTCCTGTTCCGTTCGCCTATATACTTTCTTTTCA

    GGCACCACTGTTCACCTTAGCCTCCATCACCACTGACATCATCGTTACTGAACATGCTAA
301 ---------+---------+---------+---------+---------+---------+  360
    CCGTGGTGACAAGTGGAATCGGAGGTAGTGGTGACTGTAGTAGCAATGACTTGTACGATT

    TCAGGCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGT
361 ---------+---------+---------+---------+---------+---------+  420
    AGTCCGGTTCCTCTGAGACATACTCTAATTGTATCTGTTTATGCTGCCGTAGCAGACACA

    CGGCGGAGATGGTATGTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGCAGAGGAG
421 ---------+---------+---------+---------+---------+---------+  480
    GCCGCCTCTACCATACAAGTCGCTCCACGACGTGCCAGACTAACCCTCCTGCGTCTCCTC

    CGCCGGGGTCGACCAGAACCACCCCCGGGCTGTGCTGGTCCCCAGTAGCCTCCGGATTGG
481 ---------+---------+---------+---------+---------+---------+  540
    GCGGCCCCAGCTGGTCTTGGTGGGGGCCCGACACGACCAGGGGTCATCGGAGGCCTAACC

    AATCATTCCCGCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCACCAGCGACGC
541 ---------+---------+---------+---------+---------+---------+  600
    TTAGTAAGGGCGTCCCAGTTGCCTGACGCACACAATGAGGTGGCACCCGTGGTCGCTGCG

    AGAAACCTCGGCGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTGTCCTCAGT
601 ---------+---------+---------+---------+---------+---------+  660
    TCTTTGGAGCCGCGACGTATAGCAACAACCCCTGAGCGACCGGTACCTACACAGGAGTCA

    CCACCACAACAGCACACTCCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGG
661 ---------+---------+---------+---------+---------+---------+  720
    GGTGGTGTTGTCGTGTGAGGAAGCGATGAGGCACAGGGACGACCCGATGCCGAAGATGCC

    GGACATCATCAAGGACAGTGAGAAGAAACGGTGGTTGGGTCTTGCCAGATACGACTTTTC
721 ---------+---------+---------+---------+---------+---------+  780
    CCTGTAGTAGTTCCTGTCACTCTTCTTTGCCACCAACCCAGAACGGTCTATGCTGAAAAG
```

Figure 7 cont'd

```
      AGGTTTAAAGACCTTCCTCTCCCACCACTGCTATGAAGGGACAGTGTCCTTCCTCCCTGC
781   ---------+---------+---------+---------+---------+---------+   840
      TCCAAATTTCTGGAAGGAGAGGGTGGTGACGATACTTCCCTGTCACAGGAAGGAGGGACG

      ACAACACACGGTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGGCAGGATGCTTTGTTTG
841   ---------+---------+---------+---------+---------+---------+   900
      TGTTGTGTGCCACCCTAGAGGTTCCCTATCCTTCGGGACGGCCCGTCCTACGAAACAAAC

      CAGGCAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAGAAAGCACTGTATGGTTTGGAAGC
901   ---------+---------+---------+---------+---------+---------+   960
      GTCCGTTTCGTTCGTCGTCGACCTCCTCCTCGTCTTCTTTCGTGACATACCAAACCTTCG

      TGCGGAGGACGTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAC
961   ---------+---------+---------+---------+---------+---------+  1020
      ACGCCTCCTGCACCTCCTCACCGTTCAGCAGACACCCTTCAAAGACCGGTAGTTACGGTG

      AAACATGTCCTGTGCTTGTCGCCGGAGCCCCAGGGGCCTCTCCCCGGCTGCCCACTTGGG
1021  ---------+---------+---------+---------+---------+---------+  1080
      TTTGTACAGGACACGAACAGCGGCCTCGGGGTCCCCGGAGAGGGGCCGACGGGTGAACCC

      AGACGGGTCTTCTGACCTCATCCTCATCCGGAAATGCTCCAGGTTCAATTTTCTGAGATT
1081  ---------+---------+---------+---------+---------+---------+  1140
      TCTGCCCAGAAGACTGGAGTAGGAGTAGGCCTTTACGAGGTCCAAGTTAAAAGACTCTAA

      TCTCATCAGGCACACCAACCAGCAGGACCAGTTTGACTTCACTTTTGTTGAAGTTTATCG
1141  ---------+---------+---------+---------+---------+---------+  1200
      AGAGTAGTCCGTGTGGTTGGTCGTCCTGGTCAAACTGAAGTGAAAACAACTTCAAATAGC

      CGTCAAGAAATTCCAGTTTACGTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGGA
1201  ---------+---------+---------+---------+---------+---------+  1260
      GCAGTTCTTTAAGGTCAAATGCAGCTTCGTGTACCTCCTACTCCTGTCGCTGGAGTTCCT

      GGGGGGGAAGAAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCCTGCTGCTGCACCGT
1261  ---------+---------+---------+---------+---------+---------+  1320
      CCCCCCCTTCTTCGCGAAACCCGTGTAAACGTCGTCGGTGGGGAGGACGACGACGTGGCA

      CTCCAACAGCTCCTGGAACTGCGACGGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAG
1321  ---------+---------+---------+---------+---------+---------+  1380
      GAGGTTGTCGAGGACCTTGACGCTGCCCCTCCAGGACGTGTCGGGACGGTAGCTCCAGTC

      AGTCCACTGCCAGCTGGTTCGACTCTTTGCACGAGAATTGGAAGAGAATCCGAAGCCAGA
1381  ---------+---------+---------+---------+---------+---------+  1440
      TCAGGTGACGGTCGACCAAGCTGAGAAACGTGCTCTTAACCTTCTCTTAGGCTTCGGTCT

      CTCACACAGCTGAGAAGCCGGCGTCCTGCTCTCGAACTGGGAAAGTGTGAAAACTATTTA
1441  ---------+---------+---------+---------+---------+---------+  1500
      GAGTGTGTCGACTCTTCGGCCGCAGGACGAGAGCTTGACCCTTTCACACTTTTGATAAAT

      AGATAATTATTACAGACCAATTATGTTGATATATACATTTAAATGTAGAAATTTATTTTT
1501  ---------+---------+---------+---------+---------+---------+  1560
      TCTATTAATAATGTCTGGTTAATACAACTATATATGTAAATTTACATCTTTAAATAAAAA

      GATAGTTAAATCTTGATTTTAGAAGAAAACCCTTTTGTCAACAATTTTGTGTACATATTT
1561  ---------+---------+---------+---------+---------+---------+  1620
      CTATCAATTTAGAACTAAAATCTTCTTTTGGGAAAACAGTTGTTAAAACACATGTATAAA
```

Figure 7 cont'd


```
      GGCATTTTCAGTTCTGTACGCATCTGCGGGTTGCAGCCCACGCCGCTTACTCTCAGCGGA
1621  ---------+---------+---------+---------+---------+---------+  1680
      CCGTAAAAGTCAAGACATGCGTAGACGCCCAACGTCGGGTGCGGCGAATGAGAGTCGCCT

      TGCAGCTGCTCACTTGGGGGCACTGGCCTCTTAGGTTTTAACGATGTCAACAGTGTAGTT
1681  ---------+---------+---------+---------+---------+---------+  1740
      ACGTCGACGAGTGAACCCCCGTGACCGGAGAATCCAAAATTGCTACAGTTGTCACATCAA

      TAGAAAATGGCCCGTTAGTGGCTCTATTGCAATAATGTTAGGGACATTATATGATTTCCA
1741  ---------+---------+---------+---------+---------+---------+  1800
      ATCTTTTACCGGGCAATCACCGAGATAACGTTATTACAATCCCTGTAATATACTAAAGGT

      CGCAGGTCACACCATCTGGGCCTGAGGTAGCAGTGGGTCACTTTGATCCACTTTGCAGGA
1801  ---------+---------+---------+---------+---------+---------+  1860
      GCGTCCAGTGTGGTAGACCCGGACTCCATCGTCACCCAGTGAAACTAGGTGAAACGTCCT

      CTTATTCTGTAACGGTTTGTGGCCAAGTTTTGGGAAGTGGTTGATTCTCTTTGCCTTCAT
1861  ---------+---------+---------+---------+---------+---------+  1920
      GAATAAGACATTGCCAAACACCGGTTCAAAACCCTTCACCAACTAAGAGAAACGGAAGTA

      TTCACCTTCCTCTTCGTTTACGGTTAGGACATCGCTGCTTGATCCTTACAATACTGTGCA
1921  ---------+---------+---------+---------+---------+---------+  1980
      AAGTGGAAGGAGAAGCAAATGCCAATCCTGTAGCGACGAACTAGGAATGTTATGACACGT

      ACTGCAATGCAACGTGGCCCTGCTTCAGGTGATCCGCGGGAGGGGCCTCCACGCCAGCAC
1981  ---------+---------+---------+---------+---------+---------+  2040
      TGACGTTACGTTGCACCGGGACGAAGTCCACTAGGCGCCCTCCCCGGAGGTGCGGTCGTG

      CGGGAAAGGCTGCTGGGGCCTCCACACCTGCCTCATCACGGGGGGGAAGCTACGACAATC
2041  ---------+---------+---------+---------+---------+---------+  2100
      GCCCTTTCCGACGACCCCGGAGGTGTGGACGGAGTAGTGCCCCCCCTTCGATGCTGTTAG

      CGGCTGGGAACATGACCTTGGCGTCTGTTCTGGGAACACAAATRATAARCTCTGGAARCT
2101  ---------+---------+---------+---------+---------+---------+  2160
      GCCGACCCTTGTACTGGAACCGCAGACAAGACCCTTGTGTTTAYTATTYGAGACCTTYGA

      GGCAGTGTGTAAAGCACTGGCAAGTTTGTTACTGTTAAAATGTCAAATACCAATGCTTTA
2161  ---------+---------+---------+---------+---------+---------+  2220
      CCGTCACACATTTCGTGACCGTTCAAACAATGACAATTTTACAGTTTATGGTTACGAAAT

      TATCGACGCCGAAATGCTTAACACAKCCGGGCTTGGGGGCAGTCAGGAAGAAAACTGGCC
2221  ---------+---------+---------+---------+---------+---------+  2280
      ATAGCTGCGGCTTTACGAATTGTGTMGGCCCGAACCCCCGTCAGTCCTTCTTTTGACCGG

      ATCCGTGGAGGAGGGGCCGGTCCTGGACTCCCGCAGGAYTCCTCTGATGCAGGGCCTGAA
2281  ---------+---------+---------+---------+---------+---------+  2340
      TAGGCACCTCCTCCCCGGCCAGGACCTGAGGGCGTCCTRAGGAGACTACGTCCCGGACTT

      GTCTGTACACGTGGTCCAAATTTGTCCTTGTCTTTTCTTCACACTGAGTTCTCTATATTT
2341  ---------+---------+---------+---------+---------+---------+  2400
      CAGACATGTGCACCAGGTTTAAACAGGAACAGAAAAGAAGTGTGACTCAAGAGATATAAA

      ATTGAACATCTTGTCCTTTTTAARCCAMGAAGTARTGTTAACTGCGTCTCGGATGTCTGT
2401  ---------+---------+---------+---------+---------+---------+  2460
      TAACTTGTAGAACAGGAAAAATTYGGTKCTTCATYACAATTGACGCAGAGCCTACAGACA
```

Figure 7 cont'd


```
      CTTTTGTCTCSGAARCCACRAWGGATCKCTGGTTTCCTCTCTGCAGCGCGAGGGCTCCGG
2461  ---------+---------+---------+---------+---------+---------+  2520
      GAAAACAGAGSCTTYGGTGYTWCCTAGMGACCAAAGGAGAGACGTCGCGCTCCCGAGGCC

      CGACCAGAGGATYCTYCCCGRAAGGSATTCCTGCCGCGCTCCCCCGGGCACCCCTCAATT
2521  ---------+---------+---------+---------+---------+---------+  2580
      GCTGGTCTCCTARGARGGGCYTTCCSTAAGGACGGCGCGAGGGGGCCCGTGGGGAGTTAA

      GTGTACTACCGTCCCTTGTTTAAKGGTTTGTATCCCTGCCCACSTAAGATAAATGTCTGT
2581  ---------+---------+---------+---------+---------+---------+  2640
      CACATGATGGCAGGGAACAAATTMCCAAACATAGGGACGGGTGSATTCTATTTACAGACA

      AACGGTAGTTTTGTTTGAAAATATGAGAATATGCGGCTTAAACTTTGATCTGTAAGGAGC
2641  ---------+---------+---------+---------+---------+---------+  2700
      TTGCCATCAAAACAAACTTTTATACTCTTATACGCCGAATTTGAAACTAGACATTCCTCG

      GGGGCCCGTGCCCGTTTGGAGCACGCTGTAGACMCCGTTCCTCATGCTGCCGGGTGGGTT
2701  ---------+---------+---------+---------+---------+---------+  2760
      CCCCGGGCACGGGCAAACCTCGTGCGACATCTGKGGCAAGGAGTACGACGGCCCACCCAA

      TTGCAGAAGCTCCCTTAGTGATTTCATGTTTAACAGGCAGCATCCCATTTTCAGAATTTC
2761  ---------+---------+---------+---------+---------+---------+  2820
      AACGTCTTCGAGGGAATCACTAAAGTACAAATTGTCCGTCGTAGGGTAAAAGTCTTAAAG

      CTGGCATTGATTTTATATTTTGAAGCATACAGGAAACTTCTCGTTTCCCTCGTTTAGCCC
2821  ---------+---------+---------+---------+---------+---------+  2880
      GACCGTAACTAAAATATAAAACTTCGTATGTCCTTTGAAGAGCAAAGGGAGCAAATCGGG

      CCACCCAGATCCAGGTGAAAGGGCAGCTTTAATGGTGGTTTTTATGGACCACCATTATCA
2881  ---------+---------+---------+---------+---------+---------+  2940
      GGTGGGTCTAGGTCCACTTTCCCGTCGAAATTACCACCAAAAATACCTGGTGGTAATAGT

      GAGAGCACTGTGCAAGCCAAATGGTTCCAATAATGAATGAAAATTTCTGGGTGTAAAGAG
2941  ---------+---------+---------+---------+---------+---------+  3000
      CTCTCGTGACACGTTCGGTTTACCAAGGTTATTACTTACTTTTAAAGACCCACATTTCTC

      TAAATATGCCCCTGGCTCTTTTCTACCAATGTTTGCTTCCTGGTTGGAAAGAAACCAAAG
3001  ---------+---------+---------+---------+---------+---------+  3060
      ATTTATACGGGGACCGAGAAAAGATGGTTACAAACGAAGGACCAACCTTTCTTTGGTTTC

      ATTTAAGACGGGCTGCTTCTTCCCAGACTGGCTGTGCCCTGCCTGTGGCCCAGCAACCTG
3061  ---------+---------+---------+---------+---------+---------+  3120
      TAAATTCTGCCCGACGAAGAAGGGTCTGACCGACACGGGACGGACACCGGGTCGTTGGAC

      TGCAGCCGGCAGTGTGCCTGGTGTCACGCCAGGAGGCTGTGGCTGCTGTGGGCCCTCTGG
3121  ---------+---------+---------+---------+---------+---------+  3180
      ACGTCGGCCGTCACACGGACCACAGTGCGGTCCTCCGACACCGACGACACCCGGGAGACC

      AATTGTGCTCCCTCCACAAAGTTTACCCCAAAAGGTTCTTCTAAGCCTTTATTGTCCCCT
3181  ---------+---------+---------+---------+---------+---------+  3240
      TTAACACGAGGGAGGTGTTTCAAATGGGGTTTTCCAAGAAGATTCGGAAATAACAGGGGA

      GGTAAATGTTTCCCTGGCTGGGCGCGGTGGCTCCACGCCTGTAATCCCAGCACTTTGGGA
3241  ---------+---------+---------+---------+---------+---------+  3300
      CCATTTACAAAGGGACCGACCCGCGCCACCGAGGTGCGGACATTAGGGTCGTGAAACCCT
```

```
      GGCCGAGGCGGGTGGATCCACCTAAGGTCAGGAGTTTGAGATCCAGCCTGCCCAACATGG
3301  ---------+---------+---------+---------+---------+---------+  3360
      CCGGCTCCGCCCACCTAGGTGGATTCCAGTCCTCAAACTCTAGGTCGGACGGGTTGTACC

      TGAAACCTYGTTTCTACTAAAAATACACAACTTAGCCAGTCTTGTTGGCGMACGCCTGTA
3361  ---------+---------+---------+---------+---------+---------+  3420
      ACTTTGGARCAAAGATGATTTTTATGTGTTGAATCGGTCAGAACAACCGCKTGCGGACAT

      ATSTTCAGYTACTAGGGACGCTGAGGCAGGAGAATCGTTTGAACCCAAGAAAGAGGTGGA
3421  ---------+---------+---------+---------+---------+---------+  3480
      TASAAGTCRATGATCCCTGCGACTCCGTCCTCTTAGCAAACTTGGGTTCTTTCTCCACCT

      GGTTGVGGTGAGCCAAGATTGCGCCAHTGCACTCCAGCCTGGGCAACAGAGGGAGAYTCC
3481  ---------+---------+---------+---------+---------+---------+  3540
      CCAACBCCACTCGGTTCTAACGCGGTDACGTGAGGTCGGACCCGTTGTCTCCCTCTRAGG

      ATCGCCCCCCCCCCAACAAAAAAAAAAGTTTCCCATACAYTGGCSTGCCCCAAAACCCACT
3541  ---------+---------+---------+---------+---------+---------+  3600
      TAGCGGGGGGGGGTTGTTTTTTTTTTCAAAGGGTATGTRACCGSACGGGGTTTTGGGTGA

      AACAATTTTAGCAAAACAGTCCAGGCCAAAGAGGAAGCATTTYATGTTCAATAAGAAACC
3601  ---------+---------+---------+---------+---------+---------+  3660
      TTGTTAAAATCGTTTTGTCAGGTCCGGTTTCTCCTTCGTAAARTACAAGTTATTCTTTGG

      CAGCCATTCCGCATGGCTGGTTCCTGAGTGGCTYTGGTGATACTCTCCAGCCACCTGCTG
3661  ---------+---------+---------+---------+---------+---------+  3720
      GTCGGTAAGGCGTACCGACCAAGGACTCACCGARACCACTATGAGAGGTCGGTGGACGAC

      ACATTGAGAATTTCAGACYTCGGGACTGCTGTTGCGGTACCGTGTGTYTGACACCTGCCA
3721  ---------+---------+---------+---------+---------+---------+  3780
      TGTAACTCTTAAAGTCTGRAGCCCTGACGACAACGCCATGGCACACARACTGTGGACGGT

      GCAGCCCTTTGCTATTTGCGCGCAGGATGGGGGTGACTGCCCAGACATTCCCGCTAGATA
3781  ---------+---------+---------+---------+---------+---------+  3840
      CGTCGGGAAACGATAAACGCGCGTCCTACCCCCACTGACGGGTCTGTAAGGGCGATCTAT

      GGTTTTGATTTCCGGGGCAGCCTTTCAGATGCGGCAGACATACAACACCTGTACTTTAGA
3841  ---------+---------+---------+---------+---------+---------+  3900
      CCAAAACTAAAGGCCCCGTCGGAAAGTCTACGCCGTCTGTATGTTGTGGACATGAAATCT

      GTTTTAAGGGAAAAAAAAATCAGAAGTGCTGGTTAGATAGTAAAAACTTAGGATAACTTA
3901  ---------+---------+---------+---------+---------+---------+  3960
      CAAAATTCCCTTTTTTTTTAGTCTTCACGACCAATCTATCATTTTTGAATCCTATTGAAT

      GAAAGGCTAGTTTTAGCTTCCTTTGTGGCTCCCCTGGTGCAAAACAATTAGCAGTTATGC
3961  ---------+---------+---------+---------+---------+---------+  4020
      CTTTCCGATCAAAATCGAAGGAAACACCGAGGGGACCACGTTTTGTTAATCGTCAATACG

      AATGGACCTGATTCTAGTTTATTCTAATTAAGAAGTGAGGCCGGGTTTGRACTTCGTTCC
4021  ---------+---------+---------+---------+---------+---------+  4080
      TTACCTGGACTAAGATCAAATAAGATTAATTCTTCACTCCGGCCCAAACYTGAAGCAAGG

      TGAATACAATCTTGAGTAACTGGGAAAGTCTGAGTGAAAGGATGGCCTCATTCTCTTTCT
4081  ---------+---------+---------+---------+---------+---------+  4140
      ACTTATGTTAGAACTCATTGACCCTTTCAGACTCACTTTCCTACCGGAGTAAGAGAAAGA
```

Figure 7 cont'd 17/22

```
      AATCTTGCTGGTTTCAAGATTAGAAAATGGCATTATTTGATCTGAAATGTTTGAGAARAC
4141  ---------+---------+---------+---------+---------+---------+  4200
      TTAGAACGACCAAAGTTCTAATCTTTTACCGTAATAAACTAGACTTTACAAACTCTTYTG

      ACGAATAAAGTTACTTGGGCAGAAAAAAAAA
4201  ---------+---------+---------+-  4231
      TGCTTATTTCAATGAACCCGTCTTTTTTTTT
```

Figure 8. Coding region of human PSKC cDNA.

```
    ATGGAAAAGCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCACGACGGCACCGCTGG
 71 ---------+---------+---------+---------+---------+---------+ 130
    TACCTTTTCGGAATGCGAAAATGTCAAGTGACACATTTCTCTCGTGCTGCCGTGGCGACC

    AAGTGGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAG
131 ---------+---------+---------+---------+---------+---------+ 190
    TTCACCCGCGTCCACTGAAAGACCACAGGTCTCCTCGTCGACACAGTGAACACCGACGTC

    ACCCTGCGGGAGATGCTGGAGAAGCTGACGTCCAGACCAAAGCATTTACTGGTATTTATC
191 ---------+---------+---------+---------+---------+---------+ 250
    TGGGACGCCCTCTACGACCTCTTCGACTGCAGGTCTGGTTTCGTAAATGACCATAAATAG

    AACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATATATGAAAGAAAAGTGGCACCACTG
251 ---------+---------+---------+---------+---------+---------+ 310
    TTGGGCAAACCTCCTTTTCCTGTTCCGTTCGCCTATATACTTTCTTTTCACCGTGGTGAC

    TTCACCTTAGCCTCCATCACCACTGACATCATCGTTACTGAACATGCTAATCAGGCCAAG
311 ---------+---------+---------+---------+---------+---------+ 370
    AAGTGGAATCGGAGGTAGTGGTGACTGTAGTAGCAATGACTTGTACGATTAGTCCGGTTC

    GAGACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGAT
371 ---------+---------+---------+---------+---------+---------+ 430
    CTCTGAGACATACTCTAATTGTATCTGTTTATGCTGCCGTAGCAGACACAGCCGCCTCTA

    GGTATGTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGCAGAGGAGCGCCGGGGTC
431 ---------+---------+---------+---------+---------+---------+ 490
    CCATACAAGTCGCTCCACGACGTGCCAGACTAACCCTCCTGCGTCTCCTCGCGGCCCCAG

    GACCAGAACCACCCCCGGGCTGTGCTGGTCCCCAGTAGCCTCCGGATTGGAATCATTCCC
491 ---------+---------+---------+---------+---------+---------+ 550
    CTGGTCTTGGTGGGGGCCCGACACGACCAGGGGTCATCGGAGGCCTAACCTTAGTAAGGG

    GCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCACCAGCGACGCAGAAACCTCG
551 ---------+---------+---------+---------+---------+---------+ 610
    CGTCCCAGTTGCCTGACGCACACAATGAGGTGGCACCCGTGGTCGCTGCGTCTTTGGAGC

    GCGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCACCACAAC
611 ---------+---------+---------+---------+---------+---------+ 670
    CGCGACGTATAGCAACAACCCCTGAGCGACCGGTACCTACACAGGAGTCAGGTGGTGTTG

    AGCACACTCCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC
671 ---------+---------+---------+---------+---------+---------+ 730
    TCGTGTGAGGAAGCGATGAGGCACAGGGACGACCCGATGCCGAAGATGCCCCTGTAGTAG

    AAGGACAGTGAGAAGAAACGGTGGTTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAG
731 ---------+---------+---------+---------+---------+---------+ 790
    TTCCTGTCACTCTTCTTTGCCACCAACCCAGAACGGTCTATGCTGAAAAGTCCAAATTTC

    ACCTTCCTCTCCCACCACTGCTATGAAGGGACAGTGTCCTTCCTCCCTGCACAACACACG
791 ---------+---------+---------+---------+---------+---------+ 850
    TGGAAGGAGAGGGTGGTGACGATACTTCCCTGTCACAGGAAGGAGGGACGTGTTGTGTGC
```

Figure 8 cont'd

```
      GTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGGCAGGATGCTTTGTTTGCAGGCAAAGC
 851  ---------+---------+---------+---------+---------+---------+  910
      CACCCTAGAGGTTCCCTATCCTTCGGGACGGCCCGTCCTACGAAACAAACGTCCGTTTCG

      AAGCAGCAGCTGGAGGAGGAGCAGAAGAAAGCACTGTATGGTTTGGAAGCTGCGGAGGAC
 911  ---------+---------+---------+---------+---------+---------+  970
      TTCGTCGTCGACCTCCTCCTCGTCTTCTTTCGTGACATACCAAACCTTCGACGCCTCCTG

      GTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC
 971  ---------+---------+---------+---------+---------+---------+  1030
      CACCTCCTCACCGTTCAGCAGACACCCTTCAAAGACCGGTAGTTACGGTGTTTGTACAGG

      TGTGCTTGTCGCCGGAGCCCCAGGGGCCTCTCCCCGGCTGCCCACTTGGGAGACGGGTCT
1031  ---------+---------+---------+---------+---------+---------+  1090
      ACACGAACAGCGGCCTCGGGGTCCCCGGAGAGGGGCCGACGGGTGAACCCTCTGCCCAGA

      TCTGACCTCATCCTCATCCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG
1091  ---------+---------+---------+---------+---------+---------+  1150
      AGACTGGAGTAGGAGTAGGCCTTTACGAGGTCCAAGTTAAAAGACTCTAAAGAGTAGTCC

      CACACCAACCAGCAGGACCAGTTTGACTTCACTTTTGTTGAAGTTTATCGCGTCAAGAAA
1151  ---------+---------+---------+---------+---------+---------+  1210
      GTGTGGTTGGTCGTCCTGGTCAAACTGAAGTGAAAACAACTTCAAATAGCGCAGTTCTTT

      TTCCAGTTTACGTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAG
1211  ---------+---------+---------+---------+---------+---------+  1270
      AAGGTCAAATGCAGCTTCGTGTACCTCCTACTCCTGTCGCTGGAGTTCCTCCCCCCCTTC

      AAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCCTGCTGCTGCACCGTCTCCAACAGC
1271  ---------+---------+---------+---------+---------+---------+  1330
      TTCGCGAAACCCGTGTAAACGTCGTCGGTGGGGAGGACGACGACGTGGCAGAGGTTGTCG

      TCCTGGAACTGCGACGGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGC
1331  ---------+---------+---------+---------+---------+---------+  1390
      AGGACCTTGACGCTGCCCCTCCAGGACGTGTCGGGACGGTAGCTCCAGTCTCAGGTGACG

      CAGCTGGTTCGACTCTTTGCACGAGAATTGGAAGAGAATCCGAAGCCAGACTCACACAGC
1391  ---------+---------+---------+---------+---------+---------+  1450
      GTCGACCAAGCTGAGAAACGTGCTCTTAACCTTCTCTTAGGCTTCGGTCTGAGTGTGTCG

      TGA
1451  ---  1453
      ACT
```

Figure 9. Predicted amino acid sequence of human PSKC protein.

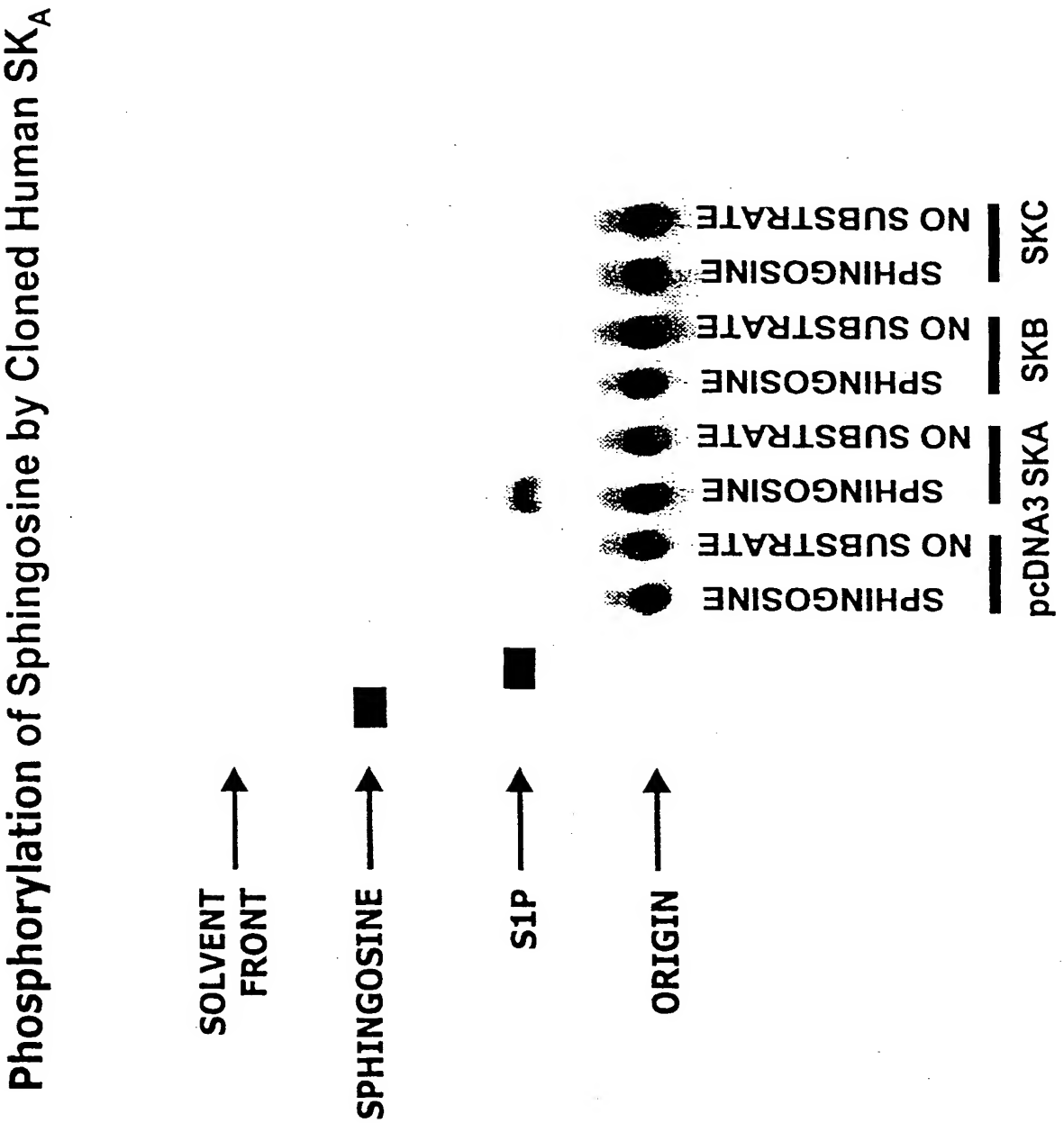
```
    MEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQLCHLWLQTLREMLEKLTSRPKHLLVFI
  1 ---------+---------+---------+---------+---------+---------+   60

    NPFGGKGQGKRIYERKVAPLFTLASITTDIIVTEHANQAKETLYEINIDKYDGIVCVGGD
 61 ---------+---------+---------+---------+---------+---------+  120

    GMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYSTVGTSDAETS
121 ---------+---------+---------+---------+---------+---------+  180

    ALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIKDSEKKRWLGLARYDFSGLK
181 ---------+---------+---------+---------+---------+---------+  240

    TFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSKQQLEEEQKKALYGLEAAED
241 ---------+---------+---------+---------+---------+---------+  300

    VEEWQVVCGKFLAINATNMSCACRRSPRGLSPAAHLGDGSSDLILIRKCSRFNFLRFLIR
301 ---------+---------+---------+---------+---------+---------+  360

    HTNQQDQFDFTFVEVYRVKKFQFTSKHMEDEDSDLKEGGKKRFGHICSSHPSCCCTVSNS
361 ---------+---------+---------+---------+---------+---------+  420

    SWNCDGEVLHSPAIEVRVHCQLVRLFARELEENPKPDSHS*
421 ---------+---------+---------+---------+-  461
```

Figure 10 21/22

Multiple alignment of novel human SK-like amino acid sequences.

```
PSKA_Human      ------------------------------------------MDPAGGPRGVLP---RP
PSKC_Human      MEKPYAFTVHCVKRARR----HRWKWAQVTFWCPEEQLCHLWLQTLREMLEKLTS---RP
PSKB_Human      MTVFFKTLRNHWKKTTAGLCLLTWGGHWLYGKHCDNLLRRAACQEAQVFGNQLIPPNAQV
                                                                  .    .

PSKA_Human      CRVLVLLNPRGGKGKALQLFRSHVQPLLAEAEISFTLMLTERRNHARELVRSEELGRWDA
PSKC_Human      KHLLVFINPFGGKGQGKRIYERKVAPLFTLASITTDIIVTEHANQAKETLYEINIDKYDG
PSKB_Human      KKATVFLNPAACKGKARTLFEKNAAPILHLSGMDVTIVKTDYEGQAKKLL--ELMENTDV
                  .  * .**   **.   ..     *.  . .  .. *.  ..*.  .    .   *

PSKA_Human      LVVMSGDGLMHEVVNGLMERPDWETAIQK--PLCILPAGSG--NALAASLNHYAGYEQVT
PSKC_Human      IVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCYSTVG
PSKB_Human      IIVAGGDGTLQEVVTGVLRRTDEATFSKI--PIGFIPLGET-----SSLSHTLFAESGN
                ..   ***   **. *.. *    .      *   .             .

PSKA_Human      NEDLLTNCTLLLCRRLLSPMNLLSLHTASGLRLFSVLSLAWGFIADVDLESEKYRRLGEM
PSKC_Human      TSDAETSALHIVVGDSLA-MDVSSVHHNSTLLRYSVSLLGYGFYGDIIKDSEKKRWLGLA
PSKB_Human      KVQHITDATLAIVKGETVPLDFLQIKGEKEQPVFAMTGLRWGSFRDAGVKVSKYWYLGPL
                     * .   .       .    ..       ... *  *   *      *   **

PSKA_Human      RFTLGTFLRLAALR--TYRGRLAYLPVGR-VGSKTPASPVVVQ---QGPVDAHLVPLEEP
PSKC_Human      RYDFSGLKTFLSHH--CYEGTVSFLPAQHTVGSPRDRKPCRAG---CFVCRQSKQQLEEE
PSKB_Human      KIKAAHFFSTLKEWPQTHQASISYTGPTERPPNEPEETPVQRPSLYRRILRRLASYWAQP
                .                .    ...                *                .

PSKA_Human      VP-----------SHWTVVPDEDFVLVLALLHSHLG--SEMFAAPMGRCAAGVMHLFYV
PSKC_Human      QKKALYGLEAAEDVEEWQVVCG-KFLAINATNMSCACRRSPRGLSPAAHLGDGSSDLILI
PSKB_Human      QD----------ALSQEVSPEVWKDVQLSTIELSITTR--NNQLDPTSKED--FLNICIE
                                        . . .  *          *  .        .

PSKA_Human      RAGVSRAMLLRLFLAMEKGRHMEYECPYLVYVPVVAFRLEPKDGKG-VFAVDGELMVRRA
PSKC_Human      RK-CSRFNFLRFLIRHTN-QQDQFDFTFVEVYRVKKFQFTSKHMEDEDSDLKEGGKKRFG
PSKB_Human      PDTISKGDFITIGSRKVR--NPKLHVEGTECLQASQCTLLIPEGAGGSFSIDSEEYEAMP
                     *.   .                                    .

PSKA_Human      VQGQVHP-------NYFWMVSGCVEPPPSWKP-------------QQMPPPEEPL
PSKC_Human      HICSSHPSCCCTVSNSSWNCDGEVLHSPAIEVRVHCQLVRLFARELEENPKPDSHS
PSKB_Human      VEVKLLPR------KLQFFCD------PRKRE-------------QMLTSPTQ--
                      *                *                .    *
```

Figure 11



SEQUENCE LISTING

<110> NPS Allelix Corp.

<120> HUMAN SPHINOSINE KINASE HOMOLOGUES

<130> p128pct3

<140> not yet issued
<141> 2000-03-02

<150> U.S. 60/122,516
<151> 1999-03-02

<160> 17

<170> PatentIn Ver. 2.0

<210> 1
<211> 1447
<212> DNA
<213> Homo sapiens

<400> 1
gaggttatgg atccagcggg cggcccccgg ggcgtgctcc cgcggccctg ccgcgtgctg 60
gtgctgctga acccgcgcgg cggcaagggc aaggccttgc agctcttccg gagtcacgtg 120
cagccccttt tggctgaggc tgaaatctcc ttcacgctga tgctcactga gcggcggaac 180
cacgcgcggg agctggtgcg gtcggaggag ctgggccgct gggacgctct ggtggtcatg 240
tctggagacg ggctgatgca cgaggtggtg aacgggctca tggagcggcc tgactgggag 300
accgccatcc agaaacccct gtgtatcctc ccagcaggct ctggcaacgc gctggcagct 360
tccttgaacc attatgctgg ctatgagcag gtcaccaatg aagacctcct gaccaactgc 420
acgctattgc tgtgccgccg gctgctgtca cccatgaacc tgctgtctct gcacacggct 480
tcggggctgc gcctcttctc tgtgctcagc ctggcctggg gcttcattgc tgatgtggac 540
ctagagagtg agaagtatcg gcgtctgggg gagatgcgct tcactctggg caccttcctg 600
cgtctggcag ccctgcgcac ctaccgcggc cgactggcct acctccctgt aggaagagtg 660
ggttccaaga cacctgcctc ccccgttgtg gtccagcagg gcccggtaga tgcacacctt 720
gtgccactgg aggagccagt gccctctcac tggacagtgg tgcccgacga ggactttgtg 780
ctagtcctgg cactgctgca ctcgcacctg ggcagtgaga tgtttgctgc acccatgggc 840
cgctgtgcag ctggcgtcat gcatctgttc tacgtgcggg cgggagtgtc tcgtgccatg 900
ctgctgcgcc tcttcctggc catggagaag ggcaggcata tggagtatga atgcccctac 960
ttggtatatg tgcccgtggt cgccttccgc ttggagccca aggatgggaa aggtgtgttt 1020
gcagtggatg gggaattgat ggttaggcga gccgtgcagg gccaggtgca cccaaactac 1080
ttctggatgg tcagcggttg cgtggagccc ccgcccagct ggaagcccca gcagatgcca 1140
ccgccagaag agcccttatg acccctgggc cgcgctgtgc cttagtgtct acttgcagga 1200
cccttcctcc ttccctaggg ctgcagggcc tgtccacagc tcctgtgggg gtggaggaga 1260
ctcctctgga gaagggtgag aaggtggagg ctatgctttg gggggacagg ccagaatgaa 1320
gtcctgggtc aggagcccag ctggctgggc ccagctgcct atgtaaggcc ttctagtttg 1380
ttctgagacc cccaccccac gaaccaaatc caaataaagt gacattccca aaaaaaaaaa 1440
aaaaaaa 1447

<210> 2
<211> 1155
<212> DNA
<213> Homo sapiens

<400> 2
atggatccag cgggcggccc ccggggcgtg ctcccgcggc cctgccgcgt gctggtgctg 60
ctgaacccgc gcggcggcaa gggcaaggcc ttgcagctct tccggagtca cgtgcagccc 120
cttttggctg aggctgaaat ctccttcacg ctgatgctca ctgagcggcg gaaccacgcg 180
cgggagctgg tgcggtcgga ggagctgggc cgctgggacg ctctggtggt catgtctgga 240
gacgggctga tgcacgaggt ggtgaacggg ctcatggagc ggcctgactg ggagaccgcc 300

```
atccagaaac ccctgtgtat cctcccagca ggctctggca acgcgctggc agcttccttg 360
aaccattatg ctggctatga gcaggtcacc aatgaagacc tcctgaccaa ctgcacgcta 420
ttgctgtgcc gccggctgct gtcacccatg aacctgctgt ctctgcacac ggcttcgggg 480
ctgcgcctct tctctgtgct cagcctggcc tggggcttca ttgctgatgt ggacctagag 540
agtgagaagt atcggcgtct gggggagatg cgcttcactc tgggcacctt cctgcgtctg 600
gcagccctgc gcacctaccg cggccgactg gcctacctcc ctgtaggaag agtgggttcc 660
aagacacctg cctcccccgt tgtggtccag cagggcccgg tagatgcaca ccttgtgcca 720
ctggaggagc cagtgccctc tcactggaca gtggtgcccg acgaggactt tgtgctagtc 780
ctggcactgc tgcactcgca cctgggcagt gagatgtttg ctgcacccat gggccgctgt 840
gcagctggcg tcatgcatct gttctacgtg cgggcgggag tgtctcgtgc catgctgctg 900
cgcctcttcc tggccatgga gaagggcagg catatggagt atgaatgccc ctacttggta 960
tatgtgcccg tggtcgcctt ccgcttggag cccaaggatg ggaaaggtgt gtttgcagtg 1020
gatggggaat tgatggttag gcgagccgtg cagggccagg tgcacccaaa ctacttctgg 1080
atggtcagcg gttgcgtgga gcccccgccc agctggaagc cccagcagat gccaccgcca 1140
gaagagccct tatga                                                  1155
```

<210> 3
<211> 384
<212> PRT
<213> Homo sapiens

<400> 3

```
Met Asp Pro Ala Gly Gly Pro Arg Gly Val Leu Pro Arg Pro Cys Arg
  1               5                  10                  15

Val Leu Val Leu Leu Asn Pro Arg Gly Gly Lys Gly Lys Ala Leu Gln
             20                  25                  30

Leu Phe Arg Ser His Val Gln Pro Leu Leu Ala Glu Ala Glu Ile Ser
         35                  40                  45

Phe Thr Leu Met Leu Thr Glu Arg Arg Asn His Ala Arg Glu Leu Val
     50                  55                  60

Arg Ser Glu Glu Leu Gly Arg Trp Asp Ala Leu Val Val Met Ser Gly
 65                  70                  75                  80

Asp Gly Leu Met His Glu Val Val Asn Gly Leu Met Glu Arg Pro Asp
                 85                  90                  95

Trp Glu Thr Ala Ile Gln Lys Pro Leu Cys Ile Leu Pro Ala Gly Ser
            100                 105                 110

Gly Asn Ala Leu Ala Ala Ser Leu Asn His Tyr Ala Gly Tyr Glu Gln
        115                 120                 125

Val Thr Asn Glu Asp Leu Leu Thr Asn Cys Thr Leu Leu Leu Cys Arg
    130                 135                 140

Arg Leu Leu Ser Pro Met Asn Leu Leu Ser Leu His Thr Ala Ser Gly
145                 150                 155                 160

Leu Arg Leu Phe Ser Val Leu Ser Leu Ala Trp Gly Phe Ile Ala Asp
                165                 170                 175

Val Asp Leu Glu Ser Glu Lys Tyr Arg Arg Leu Gly Glu Met Arg Phe
            180                 185                 190

Thr Leu Gly Thr Phe Leu Arg Leu Ala Ala Leu Arg Thr Tyr Arg Gly
        195                 200                 205
```

```
Arg Leu Ala Tyr Leu Pro Val Gly Arg Val Gly Ser Lys Thr Pro Ala
    210                 215                 220

Ser Pro Val Val Val Gln Gln Gly Pro Val Asp Ala His Leu Val Pro
225                 230                 235                 240

Leu Glu Glu Pro Val Pro Ser His Trp Thr Val Val Pro Asp Glu Asp
                245                 250                 255

Phe Val Leu Val Leu Ala Leu Leu His Ser His Leu Gly Ser Glu Met
            260                 265                 270

Phe Ala Ala Pro Met Gly Arg Cys Ala Ala Gly Val Met His Leu Phe
        275                 280                 285

Tyr Val Arg Ala Gly Val Ser Arg Ala Met Leu Leu Arg Leu Phe Leu
    290                 295                 300

Ala Met Glu Lys Gly Arg His Met Glu Tyr Glu Cys Pro Tyr Leu Val
305                 310                 315                 320

Tyr Val Pro Val Val Ala Phe Arg Leu Glu Pro Lys Asp Gly Lys Gly
                325                 330                 335

Val Phe Ala Val Asp Gly Glu Leu Met Val Arg Arg Ala Val Gln Gly
            340                 345                 350

Gln Val His Pro Asn Tyr Phe Trp Met Val Ser Gly Cys Val Glu Pro
        355                 360                 365

Pro Pro Ser Trp Lys Pro Gln Gln Met Pro Pro Pro Glu Glu Pro Leu
    370                 375                 380
```

<210> 4
<211> 2306
<212> DNA
<213> Homo sapiens

<400> 4

```
agccgcgagc tggaccagcc gtgcaaatct ctagaagatg acggtgttct ttaaaacgct 60
tcgaaatcac tggaagaaaa ctacagctgg gctctgcctg ctgacctggg gaggccattg 120
gctctatgga aaacactgtg ataacctcct aaggagagca gcctgtcaag aagctcaggt 180
gtttggcaat caactcattc ctcccaatgc acaagtgaag aaggccactg tttttctcaa 240
tcctgcagct tgcaaaggaa aagccaggac tctatttgaa aaaaatgctg ccccgatttt 300
acatttatct ggcatggatg tgactattgt taagacagat tatgagggac aagccaagaa 360
actcctggaa ctgatggaaa acacggatgt gatcattgtt gcaggaggag atgggacact 420
gcaggaggtt gttactggtg ttcttcgacg aacagatgag gctaccttca gtaagattcc 480
cattggattt atcccactgg gagagaccag tagtttgagt catacctct ttgccgaaag 540
tggaaacaaa gtccaacata ttactgatgc cacacttgcc attgtgaaag gagagacagt 600
tccacttgat ttcttgcaga tcaagggtga aaaggaacag cctgtatttg caatgaccgg 660
ccttcgatgg ggatctttca gagatgctgg cgtcaaagtt agcaagtact ggtatcttgg 720
gcctctaaaa atcaaagcag cccacttttt cagcactctt aaggagtggc ctcagactca 780
tcaagcctct atctcataca cgggacctac agagagacct cccaatgaac cagaggagac 840
ccctgtacaa aggccttctt tgtacaggag aatattacga aggcttgcgt cctactgggc 900
acaaccacag gatgcccttt cccaagaggt gagcccggag gtctggaaag atgtgcagct 960
```

gtccaccatt gaactgtcca tcacaacacg gaataatcag cttgacccga caagcaaaga 1020
agattttctg aatatctgca ttgaacctga caccatcagc aaaggagact ttataactat 1080
aggaagtcga aaggtgagaa accccaagct gcacgtggag ggcacggagt gtctccaagc 1140
cagccagtgc actttgctta tcccggaggg agcagggggc tcttttagca ttgacagtga 1200
ggagtatgaa gcgatgcctg tggaggtgaa actgctcccc aggaagctgc agttcttctg 1260
tgatcctagg aagagagaac agatgctcac aagccccacc cagtgagcag cagaagacaa 1320
gcactctgag accacacttt aggccaccgg tgggaccaaa agggaacagg tgcctcagcc 1380
atcccaacag tgtcgtcaga gggtccccag ggcattttca tggcaagtac ccctctgccc 1440
ccactccagc agtgcttccc aaagtgtgct ctgtcacctg ctttgcaatc ggcttccatt 1500
agcgcatgtt ttatttttggt gtgacggttg gccctcctaa acacggactt tcctcaggct 1560
ggttcaagac ggaaaaggac tttcttctgt tttcttccaa agtgcaacca cagtggagag 1620
cccacggtgg gcttagcctg cctaggccct tccatttctc ttctttgacc gtgctaggaa 1680
ttccaggaaa gtgcattcct gccctggtga ccttttccta tgtctaggct cctccacagg 1740
tgctgctatt ttgtgagctc cggctcctgt ttagctttta tttcagttct aacctcagtc 1800
cagaaacata tgtgaggttg tttccctctt cagccacggc tacaataccg gaaaatgcta 1860
gtttttattt attttttaa gtagtgcttc ctaaatggtt tgcatgagag ccacctgggg 1920
tacatgttga aaacttattt ggggtctacc ccaaacctaa taacccaaat ttggggatgg 1980
ggcccaggaa tatgcatttt taaaaagtca tctgcccttc ccaggtgatt ctgtaagttg 2040
tccctcaact gtacttggag aaatcgtgtt ttaaagcagt agtccacaaa gtattctgct 2100
catgtgcccc caaaagtatt ttgaaaaatc atgtataccc tcacccatct aagttgatat 2160
ctaaaatttt atctaagttg gtatctaaaa tttttcatgg gaagttaaat agttgacaaa 2220
gtatgtattt gctggtgtcg tgtaaatatt ggtattttaa aataaaaact gttacatcac 2280
taawwwaaaa aaaaaaaaaa aaaaaa 2306

<210> 5
<211> 1269
<212> DNA
<213> Homo sapiens

<400> 5
atgacggtgt tctttaaaac gcttcgaaat cactggaaga aaactacagc tgggctctgc 60
ctgctgacct ggggaggcca ttggctctat ggaaaacact gtgataacct cctaaggaga 120
gcagcctgtc aagaagctca ggtgtttggc aatcaactca ttcctcccaa tgcacaagtg 180
aagaaggcca ctgtttttct caatcctgca gcttgcaaag gaaaagccag gactctattt 240
gaaaaaaatg ctgccccgat tttacattta tctggcatgg atgtgactat tgttaagaca 300
gattatgagg gacaagccaa gaaactcctg gaactgatgg aaaacacgga tgtgatcatt 360
gttgcaggag gagatgggac actgcaggag gttgttactg gtgttcttcg acgaacagat 420
gaggctacct tcagtaagat tcccattgga tttatcccac tgggagagac cagtagtttg 480
agtcataccc tctttgccga aagtggaaac aaagtccaac atattactga tgccacactt 540
gccattgtga aaggagagac agttccactt gatttcttgc agatcaaggg tgaaaaggaa 600
cagcctgtat ttgcaatgac cggccttcga tggggatctt tcagagatgc tggcgtcaaa 660
gttagcaagt actggtatct tgggcctcta aaaatcaaag cagcccactt tttcagcact 720
cttaaggagt ggcctcagac tcatcaagcc tctatctcat acacgggacc tacagagaga 780
cctcccaatg aaccagagga gacccctgta caaaggcctt ctttgtacag gagaatatta 840
cgaaggcttg cgtcctactg ggcacaacca caggatgccc tttcccaaga ggtgagcccg 900
gaggtctgga aagatgtgca gctgtccacc attgaactgt ccatcacaac acggaataat 960
cagcttgacc cgacaagcaa agaagatttt ctgaatatct gcattgaacc tgacaccatc 1020
agcaaaggag actttataac tataggaagt cgaaaggtga gaaaccccaa gctgcacgtg 1080
gagggcacgg agtgtctcca agccagccag tgcactttgc ttatcccgga gggagcaggg 1140
ggctctttta gcattgacag tgaggagtat gaagcgatgc ctgtggaggt gaaactgctc 1200
cccaggaagc tgcagttctt ctgtgatcct aggaagagag aacagatgct cacaagcccc 1260
acccagtga 1269

<210> 6
<211> 422
<212> PRT
<213> Homo sapiens

<400> 6
Met Thr Val Phe Phe Lys Thr Leu Arg Asn His Trp Lys Lys Thr Thr

```
1               5                  10                  15

Ala Gly Leu Cys Leu Leu Thr Trp Gly Gly His Trp Leu Tyr Gly Lys
            20                  25                  30

His Cys Asp Asn Leu Leu Arg Arg Ala Ala Cys Gln Glu Ala Gln Val
        35                  40                  45

Phe Gly Asn Gln Leu Ile Pro Pro Asn Ala Gln Val Lys Lys Ala Thr
    50                  55                  60

Val Phe Leu Asn Pro Ala Ala Cys Lys Gly Lys Ala Arg Thr Leu Phe
65                  70                  75                  80

Glu Lys Asn Ala Ala Pro Ile Leu His Leu Ser Gly Met Asp Val Thr
                85                  90                  95

Ile Val Lys Thr Asp Tyr Glu Gly Gln Ala Lys Lys Leu Leu Glu Leu
            100                 105                 110

Met Glu Asn Thr Asp Val Ile Ile Val Ala Gly Gly Asp Gly Thr Leu
        115                 120                 125

Gln Glu Val Val Thr Gly Val Leu Arg Arg Thr Asp Glu Ala Thr Phe
    130                 135                 140

Ser Lys Ile Pro Ile Gly Phe Ile Pro Leu Gly Glu Thr Ser Ser Leu
145                 150                 155                 160

Ser His Thr Leu Phe Ala Glu Ser Gly Asn Lys Val Gln His Ile Thr
                165                 170                 175

Asp Ala Thr Leu Ala Ile Val Lys Gly Glu Thr Val Pro Leu Asp Phe
            180                 185                 190

Leu Gln Ile Lys Gly Glu Lys Glu Gln Pro Val Phe Ala Met Thr Gly
        195                 200                 205

Leu Arg Trp Gly Ser Phe Arg Asp Ala Gly Val Lys Val Ser Lys Tyr
    210                 215                 220

Trp Tyr Leu Gly Pro Leu Lys Ile Lys Ala Ala His Phe Phe Ser Thr
225                 230                 235                 240

Leu Lys Glu Trp Pro Gln Thr His Gln Ala Ser Ile Ser Tyr Thr Gly
                245                 250                 255

Pro Thr Glu Arg Pro Pro Asn Glu Pro Glu Glu Thr Pro Val Gln Arg
            260                 265                 270

Pro Ser Leu Tyr Arg Arg Ile Leu Arg Arg Leu Ala Ser Tyr Trp Ala
        275                 280                 285

Gln Pro Gln Asp Ala Leu Ser Gln Glu Val Ser Pro Glu Val Trp Lys
    290                 295                 300

Asp Val Gln Leu Ser Thr Ile Glu Leu Ser Ile Thr Thr Arg Asn Asn
305                 310                 315                 320

Gln Leu Asp Pro Thr Ser Lys Glu Asp Phe Leu Asn Ile Cys Ile Glu
```

```
                325                 330                 335

Pro Asp Thr Ile Ser Lys Gly Asp Phe Ile Thr Ile Gly Ser Arg Lys
            340                 345                 350

Val Arg Asn Pro Lys Leu His Val Glu Gly Thr Glu Cys Leu Gln Ala
        355                 360                 365

Ser Gln Cys Thr Leu Leu Ile Pro Glu Gly Ala Gly Gly Ser Phe Ser
    370                 375                 380

Ile Asp Ser Glu Glu Tyr Glu Ala Met Pro Val Glu Val Lys Leu Leu
385                 390                 395                 400

Pro Arg Lys Leu Gln Phe Phe Cys Asp Pro Arg Lys Arg Glu Gln Met
                405                 410                 415

Leu Thr Ser Pro Thr Gln
            420
```

<210> 7
<211> 4231
<212> DNA
<213> Homo sapiens

<400> 7

```
atctgagatc atcgccgttg aggaaacaga cgttcacggg aaacatcaag gcagtggaaa 60
atggcagaaa atggaaaagc cttacgcttt tacagttcac tgtgtaaaga gagcacgacg 120
gcaccgctgg aagtgggcgc aggtgacttt ctggtgtcca gaggagcagc tgtgtcactt 180
gtggctgcag accctgcggg agatgctgga gaagctgacg tccagaccaa agcatttact 240
ggtatttatc aacccgtttg gaggaaaagg acaaggcaag cggatatatg aaagaaaagt 300
ggcaccactg ttcaccttag cctccatcac cactgacatc atcgttactg aacatgctaa 360
tcaggccaag gagactctgt atgagattaa catagacaaa tacgacggca tcgtctgtgt 420
cggcggagat ggtatgttca gcgaggtgct gcacggtctg attgggagga cgcagaggag 480
cgccggggtc gaccagaacc acccccgggc tgtgctggtc cccagtagcc tccggattgg 540
aatcattccc gcagggtcaa cggactgcgt gtgttactcc accgtgggca ccagcgacgc 600
agaaacctcg gcgctgcata tcgttgttgg ggactcgctg gccatggatg tgtcctcagt 660
ccaccacaac agcacactcc ttcgctactc cgtgtccctg ctgggctacg gcttctacgg 720
ggacatcatc aaggacagtg agaagaaacg gtggttgggt cttgccagat acgacttttc 780
aggtttaaag accttcctct cccaccactg ctatgaaggg acagtgtcct tcctccctgc 840
acaacacacg gtgggatctc caagggatag gaagccctgc cgggcaggat gctttgtttg 900
caggcaaagc aagcagcagc tggaggagga gcagaagaaa gcactgtatg gtttggaagc 960
tgcggaggac gtggaggagt ggcaagtcgt ctgtgggaag tttctggcca tcaatgccac 1020
aaacatgtcc tgtgcttgtc gccggagccc caggggcctc tccccggctg cccacttggg 1080
agacgggtct tctgacctca tcctcatccg gaaatgctcc aggttcaatt ttctgagatt 1140
tctcatcagg cacaccaacc agcaggacca gtttgacttc acttttgttg aagtttatcg 1200
cgtcaagaaa ttccagttta cgtcgaagca catggaggat gaggacagcg acctcaagga 1260
ggggggggaag aagcgctttg ggcacatttg cagcagccac ccctcctgct gctgcaccgt 1320
ctccaacagc tcctggaact gcgacgggga ggtcctgcac agccctgcca tcgaggtcag 1380
agtccactgc cagctggttc gactctttgc acgagaattg gaagagaatc cgaagccaga 1440
ctcacacagc tgagaagccg gcgtcctgct ctcgaactgg gaaagtgtga aaactattta 1500
agataattat tacagaccaa ttatgttgat atatacattt aaatgtagaa atttatttt 1560
gatagttaaa tcttgatttt agaagaaaac ccttttgtca acaattttgt gtacatattt 1620
ggcattttca gttctgtacg catctgcggg ttgcagccca cgccgcttac tctcagcgga 1680
tgcagctgct cacttggggg cactggcctc ttaggtttta acgatgtcaa cagtgtagtt 1740
tagaaaatgg cccgttagtg gctctattgc aataatgtta gggacattat atgatttcca 1800
cgcaggtcac accatctggg cctgaggtag cagtgggtca ctttgatcca ctttgcagga 1860
cttattctgt aacggtttgt ggccaagttt tgggaagtgg ttgattctct ttgccttcat 1920
ttcaccttcc tcttcgttta cggttaggac atcgctgctt gatccttaca atactgtgca 1980
```

actgcaatgc aacgtggccc tgcttcaggt gatccgcggg aggggcctcc acgccagcac 2040
cgggaaaggc tgctggggcc tccacacctg cctcatcacg gggggggaagc tacgacaatc 2100
cggctgggaa catgaccttg gcgtctgttc tgggaacaca aatrataarc tctggaarct 2160
ggcagtgtgt aaagcactgg caagtttgtt actgttaaaa tgtcaaatac caatgcttta 2220
tatcgacgcc gaaatgctta acacakccgg gcttgggggc agtcaggaag aaaactggcc 2280
atccgtggag gaggggccgg tcctggactc ccgcaggayt cctctgatgc agggcctgaa 2340
gtctgtacac gtggtccaaa tttgtccttg tcttttcttc acactgagtt ctctatattt 2400
attgaacatc ttgtcctttt taarccamga agtartgtta actgcgtctc ggatgtctgt 2460
cttttgtctc sgaarccacr awggatckct ggtttcctct ctgcagcgcg agggctccgg 2520
cgaccagagg atyctycccg raaggsattc ctgccgcgct cccccgggca cccctcaatt 2580
gtgtactacc gtcccttgtt taakggtttg tatccctgcc cacstaagat aaatgtctgt 2640
aacggtagtt ttgtttgaaa atatgagaat atgcggctta aactttgatc tgtaaggagc 2700
ggggcccgtg cccgtttgga gcacgctgta gacmccgttc ctcatgctgc cgggtgggtt 2760
ttgcagaagc tcccttagtg atttcatgtt taacaggcag catcccattt tcagaatttc 2820
ctggcattga ttttatattt tgaagcatac aggaaacttc tcgtttccct cgtttagccc 2880
ccacccagat ccaggtgaaa gggcagcttt aatggtggtt tttatggacc accattatca 2940
gagagcactg tgcaagccaa atggttccaa taatgaatga aaatttctgg gtgtaaagag 3000
taaatatgcc cctggctctt ttctaccaat gtttgcttcc tggttggaaa gaaaccaaag 3060
atttaagacg ggctgcttct tcccagactg gctgtgccct gcctgtggcc cagcaacctg 3120
tgcagccggc agtgtgcctg gtgtcacgcc aggaggctgt ggctgctgtg ggccctctgg 3180
aattgtgctc cctccacaaa gtttacccca aaaggttctt ctaagccttt attgtccccc 3240
ggtaaatgtt tccctggctg ggcgcggtgg ctccacgcct gtaatcccag cactttggga 3300
ggccgaggcg ggtggatcca cctaaggtca ggagtttgag atccagcctg cccaacatgg 3360
tgaaacctyg tttctactaa aaatacacaa cttagccagt cttgttggcg macgcctgta 3420
atsttcagyt actagggacg ctgaggcagg agaatcgttt gaacccaaga aagaggtgga 3480
ggttgvggtg agccaagatt gcgccahtgc actccagcct gggcaacaga gggagaytcc 3540
atcgcccccc cccaacaaaa aaaaaagttt cccatacayt ggcstgcccc aaaacccact 3600
aacaatttta gcaaaacagt ccaggccaaa gaggaagcat ttyatgttca ataagaaacc 3660
cagccattcc gcatggctgg ttcctgagtg gctytggtga tactctccag ccacctgctg 3720
acattgagaa tttcagacyt cgggactgct gttgcggtac cgtgtgtytg acacctgcca 3780
gcagcccttt gctatttgcg cgcaggatgg gggtgactgc ccagacattc ccgctagata 3840
ggttttgatt tccggggcag cctttcagat gcggcagaca tacaacacct gtactttaga 3900
gttttaaggg aaaaaaaaat cagaagtgct ggttagatag taaaaactta ggataactta 3960
gaaaggctag ttttagcttc ctttgtggct cccctggtgc aaaacaatta gcagttatgc 4020
aatggacctg attctagttt attctaatta agaagtgagg ccgggtttgr acttcgttcc 4080
tgaatacaat cttgagtaac tgggaaagtc tgagtgaaag gatggcctca ttctctttct 4140
aatcttgctg gtttcaagat tagaaaatgg cattatttga tctgaaatgt ttgagaarac 4200
acgaataaag ttacttgggc agaaaaaaaa a 4231

<210> 8
<211> 1383
<212> DNA
<213> Homo sapiens

<400> 8
atggaaaagc cttacgcttt tacagttcac tgtgtaaaga gagcacgacg gcaccgctgg 60
aagtgggcgc aggtgacttt ctggtgtcca gaggagcagc tgtgtcactt gtggctgcag 120
accctgcggg agatgctgga gaagctgacg tccagaccaa agcatttact ggtatttatc 180
aacccgtttg gaggaaaagg acaaggcaag cggatatatg aaagaaaagt ggcaccactg 240
ttcaccttag cctccatcac cactgacatc atcgttactg aacatgctaa tcaggccaag 300
gagactctgt atgagattaa catagacaaa tacgacggca tcgtctgtgt cggcggagat 360
ggtatgttca gcgaggtgct gcacggtctg attgggagga cgcagaggag cgccggggtc 420
gaccagaacc acccccgggc tgtgctggtc cccagtagcc tccggattgg aatcattccc 480
gcagggtcaa cggactgcgt gtgttactcc accgtgggca ccagcgacgc agaaacctcg 540
gcgctgcata tcgttgttgg ggactcgctg gccatggatg tgtcctcagt ccaccacaac 600
agcacactcc ttcgctactc cgtgtccctg ctgggctacg gcttctacgg ggacatcatc 660
aaggacagtg agaagaaacg gtggttgggt cttgccagat acgacttttc aggtttaaag 720
accttcctct cccaccactg ctatgaaggg acagtgtcct tcctccctgc acaacacacg 780
gtgggatctc caagggatag gaagccctgc cgggcaggat gctttgtttg caggcaaagc 840
aagcagcagc tggaggagga gcagaagaaa gcactgtatg gtttggaagc tgcggaggac 900

```
gtggaggagt ggcaagtcgt ctgtgggaag tttctggcca tcaatgccac aaacatgtcc 960
tgtgcttgtc gccggagccc caggggcctc tccccggctg cccacttggg agacgggtct 1020
tctgacctca tcctcatccg gaaatgctcc aggttcaatt ttctgagatt tctcatcagg 1080
cacaccaacc agcaggacca gtttgacttc acttttgttg aagtttatcg cgtcaagaaa 1140
ttccagttta cgtcgaagca catggaggat gaggacagcg acctcaagga ggggggggaag 1200
aagcgctttg ggcacatttg cagcagccac ccctcctgct gctgcaccgt ctccaacagc 1260
tcctggaact gcgacgggga ggtcctgcac agccctgcca tcgaggtcag agtccactgc 1320
cagctggttc gactctttgc acgagaattg gaagagaatc cgaagccaga ctcacacagc 1380
tga                                                               1383
```

<210> 9
<211> 460
<212> PRT
<213> Homo sapiens

<400> 9

```
Met Glu Lys Pro Tyr Ala Phe Thr Val His Cys Val Lys Arg Ala Arg
  1               5                  10                  15

Arg His Arg Trp Lys Trp Ala Gln Val Thr Phe Trp Cys Pro Glu Glu
             20                  25                  30

Gln Leu Cys His Leu Trp Leu Gln Thr Leu Arg Glu Met Leu Glu Lys
         35                  40                  45

Leu Thr Ser Arg Pro Lys His Leu Leu Val Phe Ile Asn Pro Phe Gly
     50          53      55                  60

Gly Lys Gly Gln Gly Lys Arg Ile Tyr Glu Arg Lys Val Ala Pro Leu
 65                  70                  75                  80

Phe Thr Leu Ala Ser Ile Thr Thr Asp Ile Ile Val Thr Glu His Ala
                 85                  90  91      93      95

Asn Gln Ala Lys Glu Thr Leu Tyr Glu Ile Asn Ile Asp Lys Tyr Asp
            100                 105                 110

Gly Ile Val Cys Val Gly Gly Asp Gly Met Phe Ser Glu Val Leu His
        115                 120                 125

Gly Leu Ile Gly Arg Thr Gln Arg Ser Ala Gly Val Asp Gln Asn His
    130                 135                 140

Pro Arg Ala Val Leu Val Pro Ser Ser Leu Arg Ile Gly Ile Ile Pro
145                 150                 155                 160

Ala Gly Ser Thr Asp Cys Val Cys Tyr Ser Thr Val Gly Thr Ser Asp
                165                 170                 175

Ala Glu Thr Ser Ala Leu His Ile Val Val Gly Asp Ser Leu Ala Met
            180                 185                 190

Asp Val Ser Ser Val His His Asn Ser Thr Leu Leu Arg Tyr Ser Val
        195                 200                 205

Ser Leu Leu Gly Tyr Gly Phe Tyr Gly Asp Ile Ile Lys Asp Ser Glu
    210                 215                 220

Lys Lys Arg Trp Leu Gly Leu Ala Arg Tyr Asp Phe Ser Gly Leu Lys
225                 230                 235                 240
```

39-711 nt

```
Thr Phe Leu Ser His His Cys Tyr Glu Gly Thr Val Ser Phe Leu Pro
                245                 250                 255

Ala Gln His Thr Val Gly Ser Pro Arg Asp Arg Lys Pro Cys Arg Ala
            260                 265                 270

Gly Cys Phe Val Cys Arg Gln Ser Lys Gln Gln Leu Glu Glu Glu Gln
        275                 280                 285

Lys Lys Ala Leu Tyr Gly Leu Glu Ala Ala Glu Asp Val Glu Glu Trp
    290                 295                 300

Gln Val Val Cys Gly Lys Phe Leu Ala Ile Asn Ala Thr Asn Met Ser
305                 310                 315                 320

Cys Ala Cys Arg Arg Ser Pro Arg Gly Leu Ser Pro Ala Ala His Leu
                325                 330                 335

Gly Asp Gly Ser Ser Asp Leu Ile Leu Ile Arg Lys Cys Ser Arg Phe
            340                 345                 350

Asn Phe Leu Arg Phe Leu Ile Arg His Thr Asn Gln Gln Asp Gln Phe
        355                 360                 365

Asp Phe Thr Phe Val Glu Val Tyr Arg Val Lys Lys Phe Gln Phe Thr
    370                 375                 380

Ser Lys His Met Glu Asp Glu Asp Ser Asp Leu Lys Glu Gly Gly Lys
385                 390                 395                 400

Lys Arg Phe Gly His Ile Cys Ser Ser His Pro Ser Cys Cys Cys Thr
                405                 410                 415

Val Ser Asn Ser Ser Trp Asn Cys Asp Gly Glu Val Leu His Ser Pro
            420                 425                 430

Ala Ile Glu Val Arg Val His Cys Gln Leu Val Arg Leu Phe Ala Arg
        435                 440                 445

Glu Leu Glu Glu Asn Pro Lys Pro Asp Ser His Ser
    450                 455                 460
```

275 → 822 nt

<210> 10
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 10
aacccgcgcg gcgcaagggc aaggc 25

<210> 11
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 11
aagggcaagg ccttgcagct cttcc 25

<210> 12
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 12
caggccgctc catgagcccg ttcac 25

<210> 13
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 13
gcatcagccc gtctccagac atga 24

<210> 14
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 14
tgcaaatctc tagaagatga cggtg 25

<210> 15
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 15
tatactcaaa ctactggtct ctccaag 27

<210> 16
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 16

ttaacataga caaatacgac ggcatcg 27

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 17
acacatccat ggccagcgag tcc 23

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 September 2000 (08.09.2000)

PCT

(10) International Publication Number
WO 00/52173 A3

(51) International Patent Classification[7]: C12N 15/54, 9/12, A61K 31/70

(21) International Application Number: PCT/CA00/00223

(22) International Filing Date: 2 March 2000 (02.03.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/122,516 2 March 1999 (02.03.1999) US

(71) Applicant *(for all designated States except US)*: NPS ALLELIX CORP. [CA/CA]; 6850 Goreway Drive, Mississauga, Ontario L4V 1V7 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: MUNROE, Donald [CA/CA]; 27 Wakefield Lane, Waterdown, Ontario L0R 2N3 (CA). GUPTA, Ashwani [CA/CA]; 7031 Dunrobin Way, Mississauga, Ontario L5N 6Y4 (CA). FALZONE, Germaine, R. [CA/CA]; 34 Brady Crescent, North York, Ontario M3L 2B5 (CA).

(74) Agents: VAN ZANT, Joan, M. et al.; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montreal, Quebec H3A 2Y3 (CA).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

(88) Date of publication of the international search report:
15 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/52173 A3

(54) Title: CLONED HUMAN SPHINGOSINE KINASE HOMOLOGUES

(57) Abstract: The present invention provides newly identified and isolated polynucleotides and their polypeptides and their uses and in particular to newly identified and isolated polynucleotides and polypeptides of the sphingosine kinase family. Three isolated polynucleotides and polypeptides for three human SK homologues are described: SKA, SKB and SKC.

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/CA 00/00223

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/54 C12N9/12 A61K31/70

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOHAMA, T. ET AL.: "Molecular Cloning and Functional Characterization of Murine Sphingosine Kinase" J. BIOL. CHEM., vol. 273, no. 37, 11 September 1998 (1998-09-11), pages 23722-23728, XP002150781 figure 1	1,2,4,6, 8-12
A	BANNO, Y. ET AL.: "Evidence for the presence of multiple forms of Sph kinase in human platelets" BIOCHEM. J., vol. 335, 1998, pages 301-304, XP002150782 page 304, column 1, paragraph 2	1-14

-/—

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 October 2000	07/11/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Mata Vicente, T.

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int tional Application No
PCT/CA 00/00223

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
T	PITSON, S. M. ET AL.: "Human sphingosine kinase:purification, molecular cloning and characterization of the native and recombinant enzymes" BIOCHEMICAL J., vol. 350, Part 2, 1 September 2000 (2000-09-01), pages 429-441, XP000946990 page 430, column 2, paragraph 5	

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)